CHARACTERIZATION OF THE OPTIMAL PLANS FOR THE MONGE-KANTOROVICH TRANSPORT PROBLEM

CHRISTIAN LÉONARD

Abstract. We present a general method, based on conjugate duality, for solving a convex minimization problem without assuming unnecessary topological restrictions on the constraint set. It leads to dual equalities and characterizations of the minimizers without constraint qualification.
As an example of application, the Monge-Kantorovich optimal transport problem is solved in great detail. In particular, the optimal transport plans are characterized without restriction. This characterization improves the already existing literature on the subject.

Contents

1. Introduction

Although the title highlights Monge-Kantorovich optimal transport problem, the aim of this paper is twofold.

- First, one presents an *"extended" saddle-point method* for solving a convex minimization problem: It is shown how to implement the standard saddle-point method in such a way that topological restrictions on the constraint sets (the so-called constraint qualifications) may essentially be removed. Of course, so doing one has to pay the price of solving an arising new problem. Namely, one has to compute the extension of some function; this may be a rather difficult task in some situations, but it will be immediate in the Monge-Kantorovich case. This method is based on conjugate duality as developed by R.T. Rockafellar in [7]. Dual equalities and characterizations of the minimizers are obtained without constraint qualification.
- Then, these "extended" saddle-point abstract results are applied to the Monge-Kantorovich optimal transport problem. In particular, the optimal plans are characterized without any restriction. This characterization improves the already existing literature on the subject.

Date: July 2006.
2000 *Mathematics Subject Classification.* 46N10, 49J45, 28A35.
Key words and phrases. Convex optimization, saddle-point, conjugate duality, optimal transport.

Other applications of the extended saddle-point method are investigated by the author in [4] in connection with entropy minimization.

The Monge-Kantorovich transport problem. Let us take A and B two Polish (separable complete metric) spaces furnished with their respective Borel σ-fields, a lower semicontinuous (cost) function $c : A\times B \to [0,\infty]$ which may take infinite values and two probability measures $\mu \in \mathcal{P}_A$ and $\nu \in \mathcal{P}_B$ on A and B. We denote $\mathcal{P}_A, \mathcal{P}_B$ and $\mathcal{P}_{AB}$ the sets of all Borel probability measures on A, B and $A\times B$. The Monge-Kantorovich problem is

$$\textit{minimize } \pi \in \mathcal{P}_{AB} \mapsto \int_{A\times B} c(a,b)\,\pi(dadb) \textit{ subject to } \pi \in P(\mu,\nu) \qquad (MK)$$

where $P(\mu,\nu)$ is the set of all $\pi \in \mathcal{P}_{AB}$ with prescribed marginals $\pi_A = \mu$ on A and $\pi_B = \nu$ on B. Note that c is measurable since it is lower semicontinuous and the integral $\int_{A\times B} c\,d\pi \in [0,\infty]$ is well-defined since $c \geq 0$.
For a general account on this active field of research, see the books of S. Rachev and L. Rüschendorf [6] and C. Villani [10, 11].

Definition 1.1 (Optimal plan). *One says that $\pi \in P(\mu,\nu)$ is an* optimal plan *if it minimizes $\gamma \mapsto \int_{A\times B} c\,d\gamma$ on $P(\mu,\nu)$* and $\int_{A\times B} c\,d\pi < \infty$.

It is well-known that there exists at least an optimal plan if and only if there exists some $\pi^o \in P(\mu,\nu)$ such that $\int_{A\times B} c\,d\pi^o < \infty$; this will be recovered at Theorem 3.2. Definition 1.1 throws away the uninteresting case where $\int_{A\times B} c\,d\pi = \infty$ for all $\pi \in P(\mu,\nu)$. Note also that, since Monge-Kantorovich problem is not a strictly convex problem, infinitely many optimal plans may exist.

Already existing optimality criteria. Some usual criteria are expressed in terms of cyclical c-monotonicity.

Definition 1.2 (Cyclically c-monotone plan). *A subset $\Gamma \subset A\times B$ is said to be* cyclically c-monotone *if for any integer $n \geq 1$ and any family $(a_1,b_1),\dots,(a_n,b_n)$ of points in Γ, $\sum_{i=1}^n c(a_i,b_i) \leq \sum_{i=1}^n c(a_i,b_{i+1})$ with the convention $b_{n+1} = b_1$.*
A probability measure $\pi \in \mathcal{P}_{AB}$ is said to be cyclically c-monotone if it is concentrated on a measurable cyclically c-monotone set Γ, i.e. $\pi(\Gamma) = 1$.

This notion goes back to the seminal paper [8] by L. Rüschendorf where the standard cyclical monotonicity of convex functions introduced by Rockafellar has been extended in view of solving Monge-Kantorovich problem.
While completing this paper, the author has been informed of the recent work [9] by W. Schachermayer and J. Teichman who have improved previous characterization criteria in several directions. The following definition introduced in [9] is useful to state [9]'s results in a concise way.

Definition 1.3 (Strongly c-monotone plan). *A transport plan $\pi \in P(\mu,\nu)$ is called* strongly c-monotone *if there exist two measurable functions φ and ψ on A and B taking their values in $[-\infty,+\infty)$ such that*

$$\begin{cases} \varphi\oplus\psi \leq c & \textit{everywhere} \\ \varphi\oplus\psi = c & \pi\textit{-almost everywhere.} \end{cases} \tag{1.4}$$

Here and below, we denote $\varphi\oplus\psi(a,b) = \varphi(a)+\psi(b)$.
One easily shows that a strongly c-monotone plan is cyclically c-monotone.
The main results of [9] are collected in the next two theorems.

Theorem 1.5 ([9])**.** *Let c be a lower semicontinuous nonnegative* finitely-valued *function. If there exists some $\pi^o \in P(\mu,\nu)$ such that $\int_{A\times B} c\,d\pi^o < \infty$, then for any $\pi \in P(\mu,\nu)$, the following three statements are equivalent:*

(i) π is an optimal plan;
(ii) π is cyclically c-monotone;
(iii) π is strongly c-monotone.

This result significantly improves an already existing criterion (see [11], Chapter 5) where the same conclusion holds with a finitely-valued function c under the following constraint qualification: There exist two nonnegative measurable functions c_A and c_B on A and B such that

$$c \le c_A \oplus c_B, \int_A c_A\,d\mu < \infty \text{ and } \int_B c_B\,d\nu < \infty. \tag{1.6}$$

Note that (1.6) implies that $\int_{A\times B} c\,d\pi < \infty$ for *all* $\pi \in P(\mu,\nu)$. It also improves a result of L. Ambrosio and A. Pratelli [1] who have shown that, when c is finitely-valued and under the moment condition

$$\begin{array}{l} \mu\left(\left\{a \in A; \int_B c(a,b)\,\nu(db) < \infty\right\}\right) > 0 \\ \nu\left(\left\{b \in B; \int_A c(a,b)\,\mu(da) < \infty\right\}\right) > 0 \end{array} \tag{1.7}$$

which is weaker than (1.6), any cyclically c-monotone π in $P(\mu,\nu)$ is both an optimal and a strongly c-monotone plan. For (1.7) to hold, it is enough that $\int_{A\times B} c\,d\mu\otimes\nu < \infty$. It is also proved in [1] that the functions φ and ψ in (1.4) can be taken such that $\varphi \in L_1(A,\mu)$ and $\psi \in L_1(B,\nu)$.

The next result is concerned with cost functions c which may take infinite values.

Theorem 1.8 ([1, 9])**.** *Let c be a lower semicontinuous $[0,\infty]$-valued function.*

(a) Any optimal plan is cyclically c-monotone.
(b) If

$$\mu\otimes\nu(\{c < \infty\}) = 1, \tag{1.9}$$

then any optimal plan is strongly c-monotone.
(c) If there exists some $\pi^o \in P(\mu,\nu)$ such that $\int_{A\times B} c\,d\pi^o < \infty$, then any strongly c-monotone plan in $P(\mu,\nu)$ is an optimal plan.

Statement (a) is proved in [1], while statements (b) and (c) are taken from [9].

Examples 1.10.

(1) An interesting example of a cyclically c-monotone plan which is not optimal is exhibited in [1], in a situation where c takes infinite values and an optimal plan exists. This is in contrast with Theorem 1.5 and emphasizes that cyclical c-monotonicity isn't the right notion to consider in the general case.
(2) Take $A = B = [0,1]$, $\mu(da) = da$, $\nu(db) = db$ the Lebesgue measure on $[0,1]$ and $c(a,b) = 0$ if $a = b$ and $+\infty$ otherwise. Condition (1.9) is restrictive enough to rule this basic situation out. In the present paper, this restriction is removed.

A new optimality criterion. Our main results about the optimal plans are Theorems 3.3 and 3.5. Next theorem sums them up.

Theorem 1.11. *Let c be a lower semicontinuous $[0,\infty]$-valued function and let $\pi \in P(\mu,\nu)$ satisfy $\int_{A\times B} c\,d\pi < \infty$.*

(a) π is an optimal plan if and only if there exist two finitely-valued functions $\varphi \in \mathbb{R}^A$ and $\psi \in \mathbb{R}^B$ such that

$$\begin{cases} \varphi \oplus \psi \le c & \text{everywhere and} \\ \varphi \oplus \psi = c & \pi\text{-almost everywhere.} \end{cases} \tag{1.12}$$

(b) If π is an optimal plan, there exist two finitely-valued functions $\varphi \in \mathbb{R}^A$ and $\psi \in \mathbb{R}^B$ such that $\varphi \in L_1(A, \mu)$, $\psi \in L_1(B, \nu)$,

$$\begin{cases} |\varphi \oplus \psi| & \le & c & \text{everywhere and} \\ \varphi \oplus \psi & = & c & \text{on } \operatorname{supp} \pi \bigcap \{c < \infty\}. \end{cases} \tag{1.13}$$

Remarks 1.14. These results improve previous literature on the subject in several aspects.

a. No restriction is imposed on c, μ and ν. In particular, (1.9) is removed.
b. For the optimality criterion (a), the so-called Kantorovich potentials φ and ψ are finitely-valued and are not required to be a priori measurable. This is in contrast with the definition of strongly c-monotone plans.
c. The analogue of (b) is usually stated as follows: If π is an optimal plan, there exist two $[-\infty, \infty)$-valued functions $\varphi \in L_1(A, \mu)$ and $\psi \in L_1(B, \nu)$ such that (1.12) holds, even in the case where c is required to be finite. The improvements carried by (1.13) are:
   - The equality $\varphi \oplus \psi = c$ holds on $\operatorname{supp} \pi \cap \{c < \infty\}$ rather than only π-almost everywhere;
   - The Kantorovich potentials φ and ψ are finitely-valued;
   - We obtain $|\varphi \oplus \psi| \le c$ rather than $\varphi \oplus \psi \le c$.

As an immediate consequence of Theorem 1.11, we obtain the following

Corollary 1.15. *Any $\pi \in P(\mu, \nu)$ satisfying $\int_{A \times B} c \, d\pi < \infty$ is an optimal plan if and only if it is strongly c-monotone.*

But the sufficient condition of Theorem 1.11 is weaker than the strong c-monotonicity, while its necessary condition is stronger.

Finally, let us indicate why considering cost functions c possibly achieving the value $+\infty$ is a significant extension. In the finite-valued case, the domain of c is the closed rectangle $A \times B$. If one wants to forbid transporting mass from A to B outside some closed subset $\mathcal{S}$ of $A \times B$ and only consider the finitely-valued lower semicontinuous cost function $\tilde{c}$ on $\mathcal{S}$, simply consider the extended cost function c on $A \times B$ which matches with $\tilde{c}$ on $\mathcal{S}$ and is $+\infty$ outside. In this case, c has a closed effective domain. But there are also lower semicontinuous functions c whose domain is an increasing union of closed subsets.

An abstract convex problem and related questions. Monge-Kantorovich problem is a particular instance of an abstract convex minimization problem which we present now.

Let $\mathcal{U}$ be a vector space, $\mathcal{L} = \mathcal{U}^*$ its algebraic dual space, Φ a $(-\infty, +\infty]$-valued convex function on $\mathcal{U}$ and Φ^* its convex conjugate for the duality $\langle \mathcal{U}, \mathcal{L} \rangle$. Let $\mathcal{Y}$ be another vector space, $\mathcal{X} = \mathcal{Y}^*$ its algebraic dual space and $T : \mathcal{L} \to \mathcal{X}$ is a linear operator. We consider the convex minimization problem

$$\text{minimize } \Phi^*(\ell) \text{ subject to } T\ell \in C, \ \ell \in \mathcal{L} \tag{P}$$

where C is a convex subset of $\mathcal{X}$. As is well known, Fenchel's duality leads to the dual problem

$$\text{maximize } \inf_{x \in C} \langle y, x \rangle - \Phi(T^* y), y \in \mathcal{Y} \tag{D}$$

where T^* is the adjoint of T.

What about Monge-Kantorovich problem? We denote C_A, C_B and C_{AB} the spaces of all continuous bounded functions on A, B and $A\times B$; C_A^*, C_B^* and C_{AB}^* are their algebraic dual spaces. Taking $\mathcal{L} = C_{AB}^*$ the algebraic dual of $\mathcal{U} = C_{AB}$, T will be the marginal operator $T\ell = (\ell_A, \ell_B) \in \mathcal{X} := C_A^* \times C_B^*$ which in restriction to those ℓ's in $\mathcal{L}$ which are probability measures gives the marginals ℓ_A on A and ℓ_B on B and C will simply be $\{(\mu,\nu)\}$. Choosing $\Phi(\varphi,\psi) = 0$ if $\varphi\oplus\psi \le c$ and $\Phi(\varphi,\psi) = +\infty$ otherwise, will lead us to Monge-Kantorovich problem.

The usual questions related to (P) and (D) are

- the dual equality: Does $\inf(P) = \sup(D)$ hold?
- the primal attainment: Does there exist a solution $\bar{\ell}$ to (P)? What about the minimizing sequences, if any?
- the dual attainment: Does there exist a solution $\bar{y}$ to (D)?
- the representation of the primal solutions: Find an identity of the type: $\bar{\ell} \in \partial\Phi(T^*\bar{y})$.

We are going to answer them in terms of some *extension* $\bar{\Phi}$ of Φ under the weak assumption

$$T^{-1}(C) \cap \operatorname{diffdom} \Phi^* \neq \emptyset \tag{1.16}$$

where

$$\operatorname{diffdom} \Phi^* = \{\ell \in \mathcal{L}; \partial_{\mathcal{L}^*}\Phi^*(\ell) \neq \emptyset\}$$

is the subset of all vectors in $\mathcal{L}$ at which Φ^* admits a nonempty subdifferential with respect to the algebraic dual pairing $\langle \mathcal{L}, \mathcal{L}^*\rangle$ where $\mathcal{L}^*$ is the algebraic dual space of $\mathcal{L}$. Note that by the geometric version of Hahn-Banach theorem, the intrinsic core of the effective domain of the objective function Φ^* : $\operatorname{icordom}\Phi^*$, is included in $\operatorname{diffdom}\Phi^*$. Hence, a useful criterion to get (1.16) is

$$T^{-1}(C) \cap \operatorname{icordom} \Phi^* \neq \emptyset. \tag{1.17}$$

The drawback of such a general approach is that one has to compute the extension $\bar{\Phi}$. In specific examples, this might be a difficult task. The extension $\bar{\Phi}$ is made precise at Section 3 for Monge-Kantorovich problem. Another important example of application of our general results is the problem of minimizing an entropy functional under a convex constraint. This is worked out by the author in [4] with probabilistic applications in mind; it is based on the explicit expression of the corresponding function $\bar{\Phi}$.

The restriction (1.17) seems very weak since $\operatorname{icordom}\Phi^*$ is the notion of interior which gives the largest possible set. As $T^{-1}(C)\cap \operatorname{dom}\Phi^* = \emptyset$ implies that (P) has no solution, the only case where the problem remains open when $\operatorname{icordom}\Phi^*$ is nonempty is the situation where $T^{-1}(C)$ and $\operatorname{dom}\Phi^*$ are tangent to each other. This is used in [4] to obtain general results for convex integral functionals.

Nevertheless, the Monge-Kantorovich optimal transport problem provides an interesting case where the constraints never stand in $\operatorname{icordom}\Phi^*$ (see Remark 3.18) so that (1.17) is useless and (1.16) is the right assumption to be used.

The strategy. A usual way to prove the dual attainment and obtain some representation of the primal solutions is to require that the constraint is qualified: a property which allows to separate the convex constraint set $T^{-1}(C)$ and the level sets of the objective function. The strategy of this article is different: one chooses ad hoc topologies so that the level sets have nonempty interiors. This also allows to apply Hahn-Banach theorem, but this time the constraint set is not required to be qualified. We take the rule not to introduce arbitrary topological assumptions since (P) is expressed without any topological notion. Because of the convexity of the problem, one takes advantage of geometric easy

properties: the topologies to be considered later are associated with seminorms which are gauges of level sets of the convex functions Φ and Φ^*. They are useful tools to work with the *geometry* of (P).
It appears that when the constraints are infinite-dimensional one can choose several different spaces $\mathcal{Y}$ without modifying the value and the solutions of (P). So that for a small space $\mathcal{Y}$ the dual attainment is not the rule. As a consequence, we are facing the problem of finding an *extension* of (D) which admits solutions in generic cases and such that the representation of the primal solution is $\bar{\ell} \in \partial\bar{\Phi}(T^*\bar{y})$ where $\bar{\Phi}$ is some extension of Φ.
We are going to

- use the standard saddle-point approach to convex problems based on conjugate duality as developed by Rockafellar in [7]
- with topologies which reflect some of the geometric structure of the objective function.

These made-to-measure topologies are associated with the gauges of the level sets of Φ and Φ^*.

Outline of the paper. The abstract results are stated without proof at Section 2. Their proofs are postponed to Section 4. Section 3 is devoted to the application of the abstract results to the Monge-Kantorovich optimal transport problem. Finally, basic results about convex minimization and gauge functionals are recalled in the Appendix.

Notation. Let X and Y be topological vector spaces. The algebraic dual space of X is X^*, the topological dual space of X is X'. The topology of X weakened by Y is $\sigma(X, Y)$ and one writes $\langle X, Y\rangle$ to specify that X and Y are in separating duality.
Let $f : X \to [-\infty, +\infty]$ be an extended numerical function. Its convex conjugate with respect to $\langle X, Y\rangle$ is $f^*(y) = \sup_{x\in X}\{\langle x, y\rangle - f(x)\} \in [-\infty, +\infty]$, $y \in Y$. Its subdifferential at x with respect to $\langle X, Y\rangle$ is $\partial_Y f(x) = \{y \in Y; f(x + \xi) \geq f(x) + \langle y, \xi\rangle, \forall \xi \in X\}$. If no confusion occurs, one writes $\partial f(x)$.
For each point a, ϵ_a is the Dirac measure at a.

Stop saying no, be strict. [1] The function $\sin x$ is not negative, but it is not nonnegative. It is not decreasing, but it is not nondecreasing. All this does not make much sense and is not far from being a nonsense for non-English speaking people. As a convention, we'll use the non-English way of saying that a *positif* function is a $[0, \infty)$-valued function while if it is $(0, \infty)$-valued it is also *strictly positif*. The integer part is a *croissant* (increasing in colloquial English) function and the exponential is also a *strictly croissant* function. Symmetrically, we also use the notions of *négatif* (negative in colloquial English) and strictly négatif, *décroissant* (decreasing in colloquial English) and strictly décroissant functions or sequences. To be coherent, $[0, \infty)$ and $(-\infty, 0]$ are respectively the sets of positif and négatif numbers, and $\epsilon > 0$ is also strictly positif. We keep the French words not to be mixed up with the usual way of writing mathematics in English.

2. The abstract convex minimization problem

In this section we give the statements of the results about the abstract convex minimization problem. The dual equality and the primal attainment are stated at Theorem 2.6; the dual attainment and the dual representation of the minimizers are stated at Theorems 2.9 and 2.13. Their proofs are postponed to Section 4.

[1]This is only a suggestion, not a demanding of the right to write maths differently.

2.1. **Basic diagram.** Let $\mathcal{Y}$ be a vector space and $\mathcal{X}=\mathcal{Y}^*$ its algebraic dual space. It is useful to define the constraint operator T by means of its adjoint $T^*:\mathcal{Y}\to\mathcal{L}^*$ ($\mathcal{L}^*$ is the algebraic dual space of $\mathcal{L}$), as follows. For all $\ell\in\mathcal{L}, x\in\mathcal{X}$,

$$T\ell = x \iff \forall y\in\mathcal{Y}, \langle T^*y,\ell\rangle_{\mathcal{L}^*,\mathcal{L}} = \langle y,x\rangle_{\mathcal{Y},\mathcal{X}}.$$

We shall assume that the restriction

$$T^*(\mathcal{Y})\subset\mathcal{U} \tag{2.1}$$

holds, where $\mathcal{U}$ is identified with a subspace of $\mathcal{L}^*=\mathcal{U}^{**}$. It follows that the diagram

$$\begin{array}{ccc} \Big\langle\, \mathcal{U} & , & \mathcal{L}\, \Big\rangle \\ T^*\Big\uparrow & & \Big\downarrow T \\ \Big\langle\, \mathcal{Y} & , & \mathcal{X}\, \Big\rangle \end{array} \qquad \text{(Diagram 0)}$$

is meaningful.

2.2. **Assumptions.** Let us give the list of our main hypotheses.

(H_Φ) 1- $\Phi:\mathcal{U}\to[0,+\infty]$ is convex and $\Phi(0)=0$
2- $\forall u\in\mathcal{U}, \exists\alpha>0, \Phi(\alpha u)<\infty$
3- $\forall u\in\mathcal{U}, u\neq 0, \exists t\in\mathbb{R}, \Phi(tu)>0$

(H_T) 1- $T^*(\mathcal{Y})\subset\mathcal{U}$
2- $\ker T^*=\{0\}$

(H_C) $C_1\triangleq C\cap\mathcal{X}_1$ is a convex $\sigma(\mathcal{X}_1,\mathcal{Y}_1)$-closed subset of $\mathcal{X}_1$

The definitions of the vector spaces $\mathcal{X}_1$ and $\mathcal{Y}_1$ which appear in the last assumption are stated below at Section 2.3. For the moment, let us only say that if C is convex and $\sigma(\mathcal{X},\mathcal{Y})$-closed, then (H_C) holds.

Comments about the assumptions.

- By construction, Φ^* is a convex $\sigma(\mathcal{L},\mathcal{U})$-closed function, even if Φ is not convex. Assuming the convexity of Φ is not a restriction.
- The assumption $(H_{\Phi 1})$ also expresses that Φ achieves its minimum at $u=0$ and that $\Phi(0)=0$. This is a practical normalization requirement which will allow us to build a gauge functional associated with Φ. More, $(H_{\Phi 1})$ implies that Φ^* also shares this property. Gauge functionals related to Φ^* will also appear later.
- With any convex function $\tilde\Phi$ satisfying $(H_{\Phi 2})$, one can associate a function Φ satisfying $(H_{\Phi 1})$ in the following manner. Because of $(H_{\Phi 2})$, $\tilde\Phi(0)$ is finite and there exists $\ell_o\in\mathcal{L}$ such that $\ell_o\in\partial\tilde\Phi(0)$. Then, $\Phi(u)\triangleq\tilde\Phi(u)-\langle\ell_o,u\rangle-\tilde\Phi(0)$, $u\in\mathcal{U}$, satisfies $(H_{\Phi 1})$ and $\tilde\Phi^*(\ell)=\Phi^*(\ell-\ell_o)-\tilde\Phi(0)$, $\ell\in\mathcal{L}$.
- The hypothesis $(H_{\Phi 3})$ is not a restriction. Indeed, assuming $(H_{\Phi 1})$, let us suppose that there exists a direction $u_o\neq 0$ such that $\Phi(tu_o)=0$ for all real t. Then any $\ell\in\mathcal{L}$ such that $\langle\ell,u_o\rangle\neq 0$ satisfies $\Phi^*(\ell)\geq\sup_{t\in\mathbb{R}} t\langle\ell,u_o\rangle=+\infty$ and can't be a solution to (P).
- The hypothesis (H_{T2}) isn't a restriction either: If $y_1-y_2\in\ker T^*$, we have $\langle T\ell,y_1\rangle=\langle T\ell,y_2\rangle$, for all $\ell\in\mathcal{L}$. In other words, the spaces $\mathcal{Y}$ and $\mathcal{Y}/\ker T^*$ both specify the same constraint sets $\{\ell\in\mathcal{L}; T\ell=x\}$.

The effective assumptions are the following ones.

- The specific form of the objective function Φ^* as a convex conjugate makes it a convex $\sigma(\mathcal{L},\mathcal{U})$-closed function.
- $(H_{\Phi 2})$ and (H_C) are geometric restrictions.

- (H_{T1}) is a regularity assumption on T.

2.3. **Variants of (P) and (D).** These variants are expressed below in terms of new spaces and functions. Let us first introduce them.

The norms $|\cdot|_\Phi$ and $|\cdot|_\Lambda$. Let $\Phi_\pm(u) = \max(\Phi(u), \Phi(-u))$. By $(H_{\Phi1})$ and $(H_{\Phi2})$, $\{u \in \mathcal{U}; \Phi_\pm(u) \le 1\}$ is a convex absorbing balanced set. Hence its gauge functional which is defined for all $u \in \mathcal{U}$ by $|u|_\Phi \stackrel{\triangle}{=} \inf\{\alpha > 0; \Phi_\pm(u/\alpha)) \le 1\}$ is a seminorm. Thanks to hypothesis $(H_{\Phi3})$, it is a norm.
Taking (H_{T1}) into account, one can define

$$\Lambda(y) \stackrel{\triangle}{=} \Phi(T^*y), y \in \mathcal{Y}. \tag{2.2}$$

Let $\Lambda_\pm(y) = \max(\Lambda(y), \Lambda(-y))$. The gauge functional on $\mathcal{Y}$ of the set $\{y \in \mathcal{Y}; \Lambda_\pm(y) \le 1\}$ is $|y|_\Lambda \stackrel{\triangle}{=} \inf\{\alpha > 0; \Lambda_\pm(y/\alpha) \le 1\}, y \in \mathcal{Y}$. Thanks to (H_Φ) and (H_T), it is a norm and

$$|y|_\Lambda = |T^*y|_\Phi, \quad y \in \mathcal{Y}. \tag{2.3}$$

The spaces. Let

$$\begin{aligned} &\mathcal{U}_1 \text{ be the } |\cdot|_\Phi\text{-completion of } \mathcal{U} \text{ and let} \\ &\mathcal{L}_1 \stackrel{\triangle}{=} (\mathcal{U}, |\cdot|_\Phi)' \text{ be the topological dual space of } (\mathcal{U}, |\cdot|_\Phi). \end{aligned}$$

Of course, we have $(\mathcal{U}_1, |\cdot|_\Phi)' \cong \mathcal{L}_1 \subset \mathcal{L}$ where any ℓ in $\mathcal{U}_1'$ is identified with its restriction to $\mathcal{U}$. Similarly, we introduce

$$\begin{aligned} &\mathcal{Y}_1 \text{ the } |\cdot|_\Lambda\text{-completion of } \mathcal{Y} \text{ and} \\ &\mathcal{X}_1 \stackrel{\triangle}{=} (\mathcal{Y}, |\cdot|_\Lambda)' \text{ the topological dual space of } (\mathcal{Y}, |\cdot|_\Lambda). \end{aligned}$$

We have $(\mathcal{Y}_1, |\cdot|_\Lambda)' \cong \mathcal{X}_1 \subset \mathcal{X}$ where any x in $\mathcal{Y}_1'$ is identified with its restriction to $\mathcal{Y}$. We also have to consider the algebraic dual space $\mathcal{L}_1^*$ and $\mathcal{X}_1^*$ of $\mathcal{L}_1$ and $\mathcal{X}_1$.

The adjoint operators of T. It will be proved at Lemma 4.1 that

$$T\mathcal{L}_1 \subset \mathcal{X}_1 \tag{2.4}$$

Let us denote T_1 the restriction of T to $\mathcal{L}_1 \subset \mathcal{L}$. By (2.4), we have $T_1 : \mathcal{L}_1 \to \mathcal{X}_1$. Let us define its adjoint $T_2^* : \mathcal{X}_1^* \to \mathcal{L}_1^*$ for all $\omega \in \mathcal{X}_1^*$ by:

$$\langle \ell, T_2^*\omega\rangle_{\mathcal{L}_1, \mathcal{L}_1^*} = \langle T_1\ell, \omega\rangle_{\mathcal{X}_1, \mathcal{X}_1^*}, \forall \ell \in \mathcal{L}_1.$$

This definition is meaningful, thanks to (2.4). We denote T_1^* the restriction of T_2^* to $\mathcal{Y}_1 \subset \mathcal{X}_1^*$. Of course, it is defined for any $y \in \mathcal{Y}_1$, by

$$\langle \ell, T_1^*y\rangle_{\mathcal{L}_1, \mathcal{L}_1^*} = \langle y, T\ell\rangle_{\mathcal{Y}_1, \mathcal{X}_1}, \ \forall \ell \in \mathcal{L}_1.$$

It will proved at Lemma 4.1 that

$$T_1^*\mathcal{Y}_1 \subset \mathcal{U}_1 \tag{2.5}$$

We have the inclusions $\mathcal{Y} \subset \mathcal{Y}_1 \subset \mathcal{X}_1^*$. The adjoint operators T^* and T_1^* are the restrictions of T_2^* to $\mathcal{Y}$ and $\mathcal{Y}_1$.

Some modifications of Φ *and* Λ. The convex conjugate of Φ the dual pairing $\langle\mathcal{U},\mathcal{L}\rangle$ is

$$\Phi^*(\ell) \triangleq \sup_{u\in\mathcal{U}}\{\langle u,\ell\rangle - \Phi(u)\}, \ell\in\mathcal{L}$$

We introduce the following modifications of Φ :

$$\begin{aligned}\Phi_0(u) &\triangleq \sup_{\ell\in\mathcal{L}}\{\langle u,\ell\rangle - \Phi^*(\ell)\}, u\in\mathcal{U}\\ \Phi_1(u) &\triangleq \sup_{\ell\in\mathcal{L}_1}\{\langle u,\ell\rangle - \Phi^*(\ell)\}, u\in\mathcal{U}_1\\ \Phi_2(\zeta) &\triangleq \sup_{\ell\in\mathcal{L}_1}\{\langle \ell,\zeta\rangle - \Phi^*(\ell)\}, \zeta\in\mathcal{L}_1^*.\end{aligned}$$

They are respectively $\sigma(\mathcal{U},\mathcal{L})$, $\sigma(\mathcal{U}_1,\mathcal{L}_1)$ and $\sigma(\mathcal{L}_1^*,\mathcal{L}_1)$-closed convex functions. It is immediate to see that the restriction of Φ_2 to $\mathcal{U}_1$ is Φ_1. As $\mathcal{L}_1=\mathcal{U}_1'$, Φ_1 is also the $|\cdot|_\Phi$-closed convex regularization of Φ. The function Φ_2 is the extension $\bar\Phi$ which appears in the introductory Section 1.
We also introduce

$$\begin{aligned}\Lambda_0(y) &\triangleq \Phi_0(T^*y),\ y\in\mathcal{Y}\\ \Lambda_1(y) &\triangleq \Phi_1(T_1^*y),\ y\in\mathcal{Y}_1\\ \Lambda_2(\omega) &\triangleq \Phi_2(T_2^*\omega),\ \omega\in\mathcal{X}_1^*\end{aligned}$$

which look like the definition (2.2). Note that thanks to (H_{T1}) and (2.5), the first equalities are meaningful. Because of the previous remarks, the restriction of Λ_2 to $\mathcal{Y}_1$ is Λ_1.

The optimization problems. Let Φ_0^* and Φ_1^* be the convex conjugates of Φ_0 and Φ_1 with respect to the dual pairings $\langle\mathcal{U},\mathcal{L}\rangle$ and $\langle\mathcal{U}_1,\mathcal{L}_1\rangle$:

$$\begin{aligned}\Phi_0^*(\ell) &\triangleq \sup_{u\in\mathcal{U}}\{\langle u,\ell\rangle - \Phi_0(u)\}, \ell\in\mathcal{L}\\ \Phi_1^*(\ell) &\triangleq \sup_{u\in\mathcal{U}_1}\{\langle u,\ell\rangle - \Phi_1(u)\}, \ell\in\mathcal{L}_1\end{aligned}$$

and Λ_0^*,Λ_1^* be the convex conjugates of Λ_0,Λ_1 with respect to the dual pairings $\langle\mathcal{Y},\mathcal{X}\rangle$ and $\langle\mathcal{Y}_1,\mathcal{X}_1\rangle$:

$$\begin{aligned}\Lambda_0^*(x) &\triangleq \sup_{y\in\mathcal{Y}}\{\langle y,x\rangle - \Lambda_0(y)\}, x\in\mathcal{X}\\ \Lambda_1^*(x) &\triangleq \sup_{y\in\mathcal{Y}_1}\{\langle y,x\rangle - \Lambda_1(y)\}, x\in\mathcal{X}_1\end{aligned}$$

Finally, denote

$$C_1 = C\cap\mathcal{X}_1.$$

The optimization problems to be considered are

minimize $\Phi^*(\ell)$	*subject to* $T\ell\in C,\quad \ell\in\mathcal{L}$	(P)
minimize $\Phi_1^*(\ell)$	*subject to* $T\ell\in C_1,\quad \ell\in\mathcal{L}_1$	(P_1)
minimize $\Lambda_0^*(x)$	*subject to* $x\in C_1,\quad x\in\mathcal{X}_1$	$(P_{1,\mathcal{X}})$
maximize $\inf_{x\in C}\langle y,x\rangle - \Lambda_0(y),$	$y\in\mathcal{Y}$	(D_0)
maximize $\inf_{x\in C_1}\langle y,x\rangle - \Lambda_1(y),$	$y\in\mathcal{Y}_1$	(D_1)
maximize $\inf_{x\in C_1}\langle x,\omega\rangle - \Lambda_2(\omega),$	$\omega\in\mathcal{X}_1^*$	(D_2)

2.4. **Statement of the abstract results.** We are now ready to give answers to the questions related to (P) and (D) in an abstract setting.

Theorem 2.6 (Primal attainment and dual equality)**.** *Assume that (H_Φ) and (H_T) hold.*

(a) *For all x in $\mathcal{X}$, we have the little dual equality*
$$\inf\{\Phi^*(\ell); \ell \in \mathcal{L}, T\ell = x\} = \Lambda_0^*(x) \in [0,\infty]. \tag{2.7}$$
Moreover, in restriction to $\mathcal{X}_1$, $\Lambda_0^* = \Lambda_1^*$ and Λ_1^* is $\sigma(\mathcal{X}_1,\mathcal{Y}_1)$-inf-compact.

(b) *The problems (P) and (P_1) are equivalent: they have the same solutions and* $\inf(P) = \inf(P_1) \in [0,\infty]$.

(c) *If C is convex and $\sigma(\mathcal{X},\mathcal{Y})$-closed, we have the dual equality*
$$\inf(P) = \sup(D_0) \in [0,\infty].$$

Assume that (H_Φ), (H_T) and (H_C) hold.

(d) *We have the dual equalities*
$$\inf(P) = \inf(P_1) = \sup(D_1) = \sup(D_2) = \inf_{x\in C}\Lambda_0^*(x) \in [0,\infty] \tag{2.8}$$

(e) *If in addition* $\inf(P) < \infty$, *then (P) is attained in $\mathcal{L}_1$. Moreover, any minimizing sequence for (P) has $\sigma(\mathcal{L}_1,\mathcal{U}_1)$-cluster points and every such cluster point solves (P).*

(f) *Let $\bar\ell \in \mathcal{L}_1$ be a solution to (P), then $\bar x \triangleq T\bar\ell$ is a solution to $(P_{1,\mathcal{X}})$ and* $\inf(P) = \Phi^*(\bar\ell) = \Lambda_0^*(\bar x)$.

Theorem 2.9 (Dual attainment and representation. Interior convex constraint)**.** *Assume that (H_Φ), (H_T) and (H_C) hold and also suppose that the interior constraint qualification*
$$C \cap \operatorname{icordom}\Lambda_0^* \neq \emptyset \tag{2.10}$$
is satisfied. Then, the following statements hold true.

(a) *The primal problem (P) is attained in $\mathcal{L}_1$ and the dual problem (D_2) is attained in $\mathcal{X}_1^*$*

(b) *Any $\bar\ell \in \mathcal{L}_1$ is a solution to (P) if and only if there exists $\bar\omega \in \mathcal{X}_1^*$ such that the following three statements hold*
$$\begin{cases} (1) & T\bar\ell \in C \\ (2) & \langle T\bar\ell,\bar\omega\rangle \le \langle x,\bar\omega\rangle \text{ for all } x \in C_1 \\ (3) & \bar\ell \in \partial_{\mathcal{L}_1}\Phi_2(T_2^*\bar\omega) \end{cases}$$
More, these three statements hold if and only if: $\bar\ell$ is a solution to (P), $\bar\omega$ is a solution to (D_2) and $\inf(P)$=$\sup(D_2)$.

It is well-known that the representation formula
$$\bar\ell \in \partial_{\mathcal{L}_1}\Phi_2(T_2^*\bar\omega) \tag{2.11}$$
is equivalent to Young's identity
$$\Phi^*(\bar\ell) + \Phi_2(T_2^*\bar\omega) = \langle T\bar\ell,\bar\omega\rangle. \tag{2.12}$$

(c) *Any solution $\bar\omega$ of (D_2) shares the following properties*
(1) *$\bar\omega$ stands in the $\sigma(\mathcal{X}_1^*,\mathcal{X}_1)$-closure of* $\operatorname{dom}\Lambda_1$.
(2) *$T_2^*\bar\omega$ stands in the $\sigma(\mathcal{L}_1^*,\mathcal{L}_1)$-closures of* $T_1^*(\operatorname{dom}\Lambda_1)$ *and* $\operatorname{dom}\Phi$.
(3) *For any x_o in $C \cap \operatorname{icordom}\Lambda_1^*$, $\bar\omega$ is $j_{D_{x_o}}$-upper semicontinuous and $j_{-D_{x_o}}$-lower semicontinuous at 0, where $j_{D_{x_o}}$ and $j_{-D_{x_o}}$ are the gauge functionals on $\mathcal{X}_1$ of the convex sets $D_{x_o} = \{x \in \mathcal{X}_1; \Lambda_0^*(x_o + x) \le \Lambda_0^*(x_o) + 1\}$ and $-D_{x_o}$.*

As will be seen at Section 3, the Monge-Kantorovich problem provides an important example where no constraint is interior (see Remark 3.18). In order to solve it without imposing constraint qualification, we are going to consider the more general situation (1.16) where the constraint is said to be a *subgradient constraint*. This means that $\bar{x} \in \operatorname{diffdom} \Lambda_0^*$ with

$$\begin{aligned} \operatorname{diffdom} \Lambda_0^* &= \{x \in \mathcal{X}_1; \partial_{\mathcal{X}_1^*} \Lambda_0^*(x) \neq \emptyset\} \quad \text{where} \\ \partial_{\mathcal{X}_1^*} \Lambda_0^*(x) &= \{\omega \in \mathcal{X}_1^*; \Lambda_0^*(x') \geq \Lambda_0^*(x) + \langle x' - x, \omega\rangle, \forall x' \in \mathcal{X}_1\}. \end{aligned}$$

Two new optimization problems to be considered are

$$\begin{aligned} &\textit{minimize } \Phi^*(\ell) && \textit{subject to } T\ell = \bar{x}, \quad \ell \in \mathcal{L} && (P^{\bar{x}}) \\ &\textit{maximize } \langle \bar{x}, \omega\rangle - \Lambda_2(\omega), && \omega \in \mathcal{X}_1^* && (D_2^{\bar{x}}) \end{aligned}$$

where $\bar{x} \in \mathcal{X}$. This corresponds to the simplified case where C is reduced to the single point $\bar{x}$.

Theorem 2.13 (Dual attainment and representation. Subgradient affine constraint). *Let us assume that* (H_Φ) *and* (H_T) *hold and suppose that* $\bar{x} \in \operatorname{dom} \Lambda_0^*$. *Then,* $\inf(P^{\bar{x}}) < \infty$. *If in addition,*

$$\bar{x} \in \operatorname{diffdom} \Lambda_0^*, \tag{2.14}$$

then the following statements hold true.

(a) The primal problem $(P^{\bar{x}})$ *is attained in* $\mathcal{L}_1$ *and the dual problem* $(D_2^{\bar{x}})$ *is attained in* $\mathcal{X}_1^*$.

(b) Any $\bar{\ell} \in \mathcal{L}_1$ *is a solution to* $(P^{\bar{x}})$ *if and only if* $T\bar{\ell} = \bar{x}$ *and there exists* $\bar{\omega} \in \mathcal{X}_1^*$ *such that (2.11) or equivalently (2.12) holds.*
More, this occurs if and only if: $\bar{\ell}$ *is a solution to* (P), $\bar{\omega}$ *is a solution to* $(D_2^{\bar{x}})$ *with* $\bar{x} := T\bar{\ell}$ *and* $\inf(P^{\bar{x}}) = \sup(D_2^{\bar{x}})$.

(c) Any solution $\bar{\omega}$ *of* $(D_2^{\bar{x}})$, *shares the following properties*
  (1) $\bar{\omega}$ *stands in the* $\sigma(\mathcal{X}_1^*, \mathcal{X}_1)$*-closure of* $\operatorname{dom} \Lambda_1$.
  (2) $T_2^*\bar{\omega}$ *stands in the* $\sigma(\mathcal{L}_1^*, \mathcal{L}_1)$*-closures of* $T_1^*(\operatorname{dom} \Lambda_1)$ *and* $\operatorname{dom} \Phi$.
  (3) Let $\bar{\omega}$ *be any solution of* $(D_2^{\bar{x}})$ *with* $\bar{x} \in \operatorname{icordom} \Lambda_0^*$. *Then,* $\bar{\omega}$ *is* $j_{D_{\bar{x}}}$*-upper semicontinuous and* $j_{-D_{\bar{x}}}$*-lower semicontinuous at* 0 *where* $D_{\bar{x}} = \{x \in \mathcal{X}_1; \Lambda_0^*(\bar{x} + x) \leq \Lambda_0^*(\bar{x}) + 1\}$.

3. Application to the Monge-Kantorovich optimal transport problem

We apply the results of Section 2 to the Monge-Kantorovich problem. Recall that we take A and B two Polish spaces furnished with their Borel σ-fields. Their product space $A \times B$ is endowed with the product topology and the corresponding Borel σ-field. The lower semicontinuous cost function $c : A \times B \to [0, \infty]$ may take infinite values. Let us also take two probability measures $\mu \in \mathcal{P}_A$ and $\nu \in \mathcal{P}_B$ on A and B. The Monge-Kantorovich problem is

$$\textit{minimize } \pi \in \mathcal{P}_{AB} \mapsto \int_{A\times B} c(a,b)\, \pi(dadb) \textit{ subject to } \pi \in P(\mu, \nu) \tag{MK}$$

where $P(\mu, \nu)$ is the set of all $\pi \in \mathcal{P}_{AB}$ with prescribed marginals $\pi_A = \mu$ on A and $\pi_B = \nu$ on B.

3.1. **Statement of the results.** Let us fix some notations. We denote C_A, C_B and C_{AB} the spaces of all continuous bounded functions on A, B and $A\times B$. The Kantorovich maximization problem:

$$\begin{aligned}&\textit{maximize } \int_A \varphi\, d\mu + \int_B \psi\, d\nu \textit{ for all } \varphi,\psi \textit{ such that}\\ &\varphi\in C_A, \psi\in C_B \textit{ and } \varphi\oplus\psi\le c\end{aligned} \tag{K}$$

is the basic dual problem of (MK). We also consider the following extended version of (K) :

$$\begin{aligned}&\textit{maximize } \int_A \varphi\, d\mu + \int_B \psi\, d\nu \textit{ for all } \varphi\in\mathbb{R}^A, \psi\in\mathbb{R}^B \textit{ such that}\\ &\varphi\in L_1(A,\mu), \psi\in L_1(B,\nu) \text{ and } \varphi\oplus\psi\le c \text{ everywhere on } A\times B.\end{aligned} \tag{$\overline{K}$}$$

Remark 3.1. The real-valued function $\varphi\in\mathbb{R}^A$ is defined *everywhere*, rather than μ-almost everywhere, and $\varphi\in L_1(A,\mu)$ implies that it is *μ-measurable*. This means that there exists some measurable set $N_A\subset A$ such that $\mu(N_A)=0$ and $\mathbf{1}_{N_A}\varphi$ is measurable. A similar remark holds for ψ.

The set of all probability measures π on $A\times B$ such that $\int_{A\times B} c\, d\pi<\infty$ is denoted $\mathcal{P}_c$. By Definition 1.1, an optimal plan stands in $\mathcal{P}_c$. In the next theorem, $\mathcal{P}_c$ will be endowed with the weak topology $\sigma(\mathcal{P}_c,\mathcal{C}_c)$ where $\mathcal{C}_c$ is the space of all continuous functions u on $A\times B$ such that $|u|\le k(1+c)$ for some $k\ge 0$.

Theorem 3.2 (Dual equality and primal attainment)**.**

(1) The dual equality for (MK) is

$$\inf(MK)=\sup(K)=\sup(\overline{K})\in[0,\infty].$$

(2) Assume that there exists some π^o in $P(\mu,\nu)$ such that $\int_{A\times B} c\, d\pi^o<\infty$. Then:
   (a) There is at least an optimal plan and all the optimal plans are in $\mathcal{P}_c$;
   (b) Any minimizing sequence is relatively compact for the topology $\sigma(\mathcal{P}_c,\mathcal{C}_c)$ and all its cluster points are optimal plans.

This result is well-known. The dual equality $\inf(MK)=\sup(K)=\sup(\overline{K})$ is the Kantorovich dual equality. The proof of Theorem 3.2 will be an opportunity to make precise the abstract material Φ, $\mathcal{U}$, $T\dots$ in terms of the Monge-Kantorovich problem.

Next, we state the characterization of the optimal plans without restriction.

Theorem 3.3 (Characterization of the optimal plans)**.**

(1) A probability measure $\pi\in\mathcal{P}_{AB}$ is an optimal plan if and only if there exist two finitely-valued functions $\varphi\in\mathbb{R}^A$ and $\psi\in\mathbb{R}^B$ such that

$$\begin{cases} (a) & \pi_A=\mu, \pi_B=\nu, \int_{A\times B} c\, d\pi<\infty,\\ (b) & \varphi\oplus\psi\le c \quad \textit{everywhere and}\\ (c) & \varphi\oplus\psi=c \quad \pi\textit{-almost everywhere.}\end{cases} \tag{3.4}$$

(2) Let φ and ψ be finitely-valued functions on A and B and let be $\pi\in\mathcal{P}_{AB}$.
   (a) If φ is μ-measurable and ψ is ν-measurable, the following statements are equivalent:
      - φ, ψ and π satisfy (3.4);
      - π is an optimal plan and (φ,ψ) is a solution of $(\overline{K})$.
   (b) In the general case where φ and ψ are not assumed to be measurable, consider the following statements:

(i) φ, ψ *and* π *satisfy (3.4);*
(ii) $\mathrm{ls}\,\varphi$, $\mathrm{ls}\,\psi$ *and* π *satisfy (3.4);*
(iii) π *is an optimal plan and* $(\mathrm{ls}\,\varphi, \mathrm{ls}\,\psi)$ *is a solution of* $(\overline{K})$.
Then: (i) $\Rightarrow$ *(ii)* $\Leftrightarrow$ *(iii).*

This new result improves the already existing literature on the subject. It is important to note that *the functions* φ *and* ψ *satisfying (3.4) are neither assumed to be integrable nor to be measurable.* Next theorem shows that they can be further specified.

Theorem 3.5 (More about necessary conditions)**.** *Assume that* π *is an optimal plan. Then, there exist two finitely-valued functions* $\varphi \in \mathbb{R}^A$ *and* $\psi \in \mathbb{R}^B$ *such that* $\varphi \in L_1(A, \mu)$, $\psi \in L_1(B, \nu)$ *and*

$$\begin{cases} |\varphi \oplus \psi| & \leq & c & \textit{everywhere and} \\ \varphi \oplus \psi & = & c & \textit{on } \operatorname{supp} \pi \bigcap \{c < \infty\}. \end{cases}$$

Clearly, (φ, ψ) *is a maximizer of* $(\overline{K})$.

Remarks 3.6.

a. Note that any optimal plan π is satisfies $\operatorname{supp} \pi \subset \operatorname{cl}\{c < \infty\}$.
b. Recall that π is said be to *concentrated* on the measurable set Γ if $\pi(\Gamma) = 1$. For instance, (3.4-c) is equivalent to the existence of some set Γ on which π is concentrated and $\varphi \oplus \psi = c$ on Γ. The *support* of π, denoted $\operatorname{supp} \pi$, is the closure of the union of all the sets Γ on which π is concentrated.

3.2. **Proof of Theorem 3.2.** We apply the general results of Section 2.

The operators T *and* T^**.* The algebraic dual spaces of C_A, C_B and C_{AB} are C_A^*, C_B^* and C_{AB}^*. We define the marginal operator

$$T\ell = (\ell_A, \ell_B) \in C_A^* \times C_B^*, \quad \ell \in C_{AB}^*$$

where $\langle \varphi, \ell_A \rangle = \langle \varphi \otimes 1, \ell \rangle$ and $\langle \psi, \ell_B \rangle = \langle 1 \otimes \psi, \ell \rangle$ for all $\varphi \in C_A$ and all $\psi \in C_B$.
Let us identify the operator T^*. For all $(\varphi, \psi) \in C_A \times C_B$ and all $\ell \in \mathcal{L}$, we have $\langle T^*(\varphi, \psi), \ell \rangle_{\mathcal{L}^*, \mathcal{L}} = \langle (\varphi, \psi), (\ell_A, \ell_B) \rangle = \langle \varphi, \ell_A \rangle + \langle \psi, \ell_B \rangle = \langle \varphi \oplus \psi, \ell \rangle_{\mathcal{U}, \mathcal{L}}$ where $\varphi \oplus \psi(a, b) = \varphi(a) + \psi(b)$. Hence, for each $\varphi \in C_A$ and $\psi \in C_B$,

$$T^*(\varphi, \psi) = \varphi \oplus \psi \in C_{A \times B}. \tag{3.7}$$

The problem (P)*.* Then, the Diagram 0 is built with $\mathcal{U} = C_{AB}$, $\mathcal{L} = C_{AB}^*$, $\mathcal{X} = C_A^* \times C_B^*$ and $\mathcal{Y} = C_A \times C_B$. Here and below, we denote the convex indicator function of the set X,

$$\delta_X(x) = \begin{cases} 0 & \text{if } x \in X \\ +\infty & \text{otherwise.} \end{cases}$$

Choosing $C = \{(\mu, \nu)\}$ and $\Phi(u) = \delta_{\{u \leq c\}}, u \in C_{AB}$ we get $\Phi^*(\ell) = \sup\{\langle u, \ell \rangle; u \in C_{AB}, u \leq c\}, \ell \in C_{AB}^*$ and we obtain the primal problem

$$\textit{minimize } \Phi^*(\ell) \textit{ subject to } \ell_A = \mu \textit{ and } \ell_B = \nu, \quad \ell \in C_{AB}^*. \tag{P}$$

It will be shown at Proposition 3.12 that the corresponding problem (P_1) is (MK).

The problem (D_0). Now, let's have a look at Φ_0. As $\{u \in C_{AB}; u \leq c\}$ is convex and $\sigma(C_{AB}, C_{AB}^*)$-closed, we have $\Phi_0 = \Phi$. Therefore, for each $\varphi \in C_A$ and $\psi \in C_B$,

$$\Lambda_0(\varphi,\psi) = \Lambda(\varphi,\psi) = \Phi(T^*(\varphi,\psi)) = \delta_{\{\varphi\oplus\psi\leq c\}}$$

and the dual problem is

$$\textit{maximize} \int_A \varphi\, d\mu + \int_B \psi\, d\nu \textit{ subject to } \varphi \oplus \psi \leq c, \quad \varphi \in C_A, \psi \in C_B \tag{D_0}$$

whose value is

$$\Lambda^*(\mu,\nu) = \sup\left\{\int_A \varphi\, d\mu + \int_B \psi\, d\nu; \varphi \in C_A, \psi \in C_B : \varphi \oplus \psi \leq c\right\}$$

As $\Lambda = \Lambda_0$ and $\Lambda_0^* = \Lambda_1^*$ (Theorem 2.6-a), we have: $\Lambda_0^* = \Lambda_1^* = \Lambda^*$.

The hypotheses (H). We begin with a simple remark.

Remark 3.8. One can choose $c \geq 1$ without loss of generality. Indeed, with $c \geq 0$ taking $\tilde{c} = c + 1$ one obtains $\int_{A\times B} c\, d\pi = \int_{A\times B} \tilde{c}\, d\pi - 1$ for all $\pi \in \mathcal{P}_{AB}$. Consequently, the minimization problems (MK) and $(\widetilde{MK})$ associated with c and $\tilde{c}$ share the same minimizers and their values are related by $\inf(MK) = \inf(\widetilde{MK}) - 1$.
It also follows from these considerations that our results still hold under the assumption that c is bounded below rather than c is positif.

We assume from now on that $c \geq 1$. This guarantees $(H_{\Phi 2})$. In the case where c is finitely valued, the remaining hypotheses (H) follow by (3.7) and direct inspection.

If c is infinite somewhere, then $(H_{\Phi 3})$ fails. Indeed, for any function $u \in C_{AB}$, we have $\Phi(tu) = 0$ for all real t if and only if $\{u \neq 0\} \subset A\times B \setminus \mathcal{S}$ where

$$\mathcal{S} = \mathrm{cl}\,\{c < \infty\}$$

is the closure of $\{(a,b) \in A\times B; c(a,b) < \infty\}$. The way to get rid of this problem is standard. Let $u \sim v$ be the equivalence relation on $\mathbb{R}^{A\times B}$ defined by $u_{|\mathcal{S}} = v_{|\mathcal{S}}$, i.e. u and v match on $\mathcal{S}$. The space $\mathcal{U}$ to be considered is the factor space

$$\mathcal{U} := C_{AB}/\sim$$

Clearly, if $u \sim v$ then $\Phi(u) = \Phi(v)$. Hence, it is possible to identify without loss of generality any $u \in C_{AB}$ with its equivalence class which in turn is identified with the restriction $u_{|\mathcal{S}}$ of u to $\mathcal{S}$.

The problem (P_1). Recall that $c \geq 1$ without loss of generality. Let us first identify the space $\mathcal{L}_1$. As $\Phi_\pm(u) = \delta_{\{|u|\leq c\}}$, we obtain the seminorm $|u|_\Phi = \sup|u/c| := \|u\|_c$ on C_{AB} which becomes a norm on $\mathcal{U}$,

$$\begin{aligned}\mathcal{U}_1 &= C_c := \{u_{|\mathcal{S}}; u : A\times B \to \mathbb{R}, u \text{ continuous and } |u| \leq kc \text{ for some real } k\}\\ \mathcal{L}_1 &= (C_c, \|\cdot\|_c)' = C_c'.\end{aligned}$$

Obviously, any π in $\mathcal{P}_c$ has its support included in $\mathcal{S}$ and belongs to C_c' with the dual bracket $\langle u_{|\mathcal{S}}, \pi\rangle = \int_{\mathcal{S}} u\, d\pi$, $u_{|\mathcal{S}} \in C_c$. In what follows, it will be written equivalently

- $u_{|\mathcal{S}} \in C_c$ to specify that the equivalence class of u stands in C_c and
- $u \in C_c$ to specify that the restriction $u_{|\mathcal{S}}$ of the continuous function u on $A\times B$ stands in C_c.

Clearly, the function Φ_1 is

$$\Phi_1(u) = \delta_{\{u \le c\}}, \quad u \in C_c$$

and the modified primal problem is

$$\textit{minimize } \Phi_1^*(\ell) \textit{ subject to } \ell_A = \mu \textit{ and } \ell_B = \nu, \quad \ell \in C_c' \qquad (P_1)$$

where for each $\ell \in C_c'$,

$$\Phi_1^*(\ell) = \sup\{\langle u, \ell \rangle; u \in C_c, u \le c\}.$$

Remark 3.9. Two representations of C_c'.

a. Let $C_c(\mathcal{S})$ be the space of all continuous functions w on $\mathcal{S}$ (w.r.t. the relative topology) such that $\|w\|_c = \sup_{\mathcal{S}} |w/c| < \infty$ and $C_c(\mathcal{S})'$ be the topological dual space of the normed space $(C_c(\mathcal{S}), \|\cdot\|_c)$. Let $\mathcal{E}$ be the subspace of all functions in $C_c(\mathcal{S})$ which can be continuously extended to the whole space $A \times B$. There is a one-one correspondence between C_c' and the dual space $(\mathcal{E}, \|\cdot\|_c)'$.
b. There is also a one-one correspondence between C_c' and the space of all linear forms ℓ on the space of all continuous functions on $A \times B$ such that $\operatorname{supp} \ell \subset \mathcal{S}$ (see Definition 3.10 below) and $\sup\{\langle u, \ell \rangle = \langle u_{|\mathcal{S}}, \ell \rangle; u : \|u\|_c \le 1\} < \infty$.

Definition 3.10. *For any linear form ℓ on the space of all continuous functions on $A \times B$, we define the* support *of ℓ as the subset of all $(a, b) \in A \times B$ such that for any neighborhood G of (a, b), there exists some function u in C_{AB} satisfying $\{u \neq 0\} \subset G$ and $\langle u, \ell \rangle \neq 0$. It is denoted* $\operatorname{supp} \ell$.

Definition 3.11.

(a) One says that $\ell \in C_c'$ acts as a probability measure *if there exists $\tilde{\ell} \in \mathcal{P}_{AB}$ such that $\operatorname{supp} \tilde{\ell} \subset \mathcal{S}$ and for all $u \in C_{AB}$, $\langle u_{|\mathcal{S}}, \ell \rangle = \int_{\mathcal{S}} u \, d\tilde{\ell}$. In this case, we write: $\ell \in \mathcal{P}_{\mathcal{S}}$.*

(b) One says that $\ell \in C_c'$ stands in $\mathcal{P}_c$ if there exists $\tilde{\ell} \in \mathcal{P}_c$ such that for all $u \in C_c$, $\langle u_{|\mathcal{S}}, \ell \rangle = \int_{\mathcal{S}} u \, d\tilde{\ell}$. In this case, we write: $\ell \in \mathcal{P}_c$.

Of course, if there exists $\tilde{\ell}$ satisfying (a), it belongs to $\mathcal{P}_c$ and is unique since any probability measure on a metric space is determined by its values on the continuous bounded functions. This explains why the notation $\ell \in \mathcal{P}_c$ in (b) isn't misleading.
Note also that any probability measure $\tilde{\ell} \in \mathcal{P}_c$ has a support included in $\mathcal{S}$. Since $A \times B$ is a metric space, for any $\ell \in \mathcal{P}_c$ acting as a measure, $\operatorname{supp} \ell$ in the sense of Definition 3.10 matches with the usual support of the measure $\tilde{\ell}$.

Completing the proof of Theorem 3.2. The full connection with the Monge-Kantorovich problem is given by the following Proposition 3.12. Clearly, with this proposition in hand, Theorem 3.2 directly follows from Theorem 2.6 and the obvious inequalities $\sup(K) \le \sup(\overline{K}) \le \inf(MK)$.

Proposition 3.12. *For all $\ell \in C_c'$,*

(a) $\Phi_1^*(\ell) < \infty \Rightarrow \ell \ge 0$,
(b) $\Phi_1^*(\ell) < \infty \Rightarrow \operatorname{supp} \ell \subset \mathcal{S}$,
(c) $[\ell \ge 0, \operatorname{supp} \ell \subset \mathcal{S}, \ell_A = \mu$ and $\ell_B = \nu] \Rightarrow \ell \in \mathcal{P}_{\mathcal{S}}$ *and*
(d) for all $\ell \in \mathcal{P}_{\mathcal{S}}$, $\Phi_1^*(\ell) = \int_{\mathcal{S}} c \, d\ell$.

It follows that

- $\operatorname{dom} \Phi_1^* \subset \mathcal{P}_c$ *and*
- *the problems (MK) and (P_1) share the same values and the same minimizers.*

Proof. Clearly, the last statement follows from the first part of the proposition. The proof is divided into four parts.

• Proof of (a). Suppose that $\ell \in C_c'$ isn't in the positif cone. This means that there exists $u_o \in C_c$ such that $u_o \geq 0$ and $\langle u_o, \ell \rangle < 0$. Since u_o satisfies $\lambda u_o \leq 0 \leq c$ for all $\lambda < 0$, we have $\Phi_1^*(\ell) \geq \sup_{\lambda<0}\{\langle \lambda u_o, \ell\rangle\} = +\infty$. Hence, $\Phi_1^*(\ell) < \infty$ implies that $\ell \geq 0$ and one can restrict our attention to the positif ℓ's.

• Proof of (b). Suppose ad absurdum that $\operatorname{supp}\ell \subsetneq \mathcal{S}$. Then, there exists a positif function $u_o \in C_{AB}$ such that $\{u_o > 0\} \cap \mathcal{S} = \emptyset$ and $\langle u_o, \ell\rangle > 0$. As $tu_o \leq c_{|A\times B\setminus \mathcal{S}} \equiv \infty$ for all $t > 0$, $\Phi_1^*(\ell) \geq \sup_{t>0}\{\langle tu_o, \ell\rangle\} = +\infty$.

• Proof of (c). Let us take $\ell \geq 0$ such that $\operatorname{supp}\ell \subset \mathcal{S}$, $\ell_A = \mu$ and $\ell_B = \nu$. It is clear that $\langle 1, \ell\rangle = 1$. It remains to check that for any $\ell \in C_c'$

$$[\ell \geq 0, \operatorname{supp}\ell \subset \mathcal{S}, \ell_A = \mu \textit{ and } \ell_B = \nu] \Rightarrow \ell \text{ is } \sigma\text{-additive}, \tag{3.13}$$

rather than only additive. Since $A\times B$ is a metric space, one can apply an extension of the construction of Daniell's integrals ([5], Proposition II.7.2) to see that ℓ acts as a measure if and only if for any décroissant sequence (u_n) of continuous functions such that $0 \leq u_n \leq 1$ for all n and $\lim_{n\to\infty} u_n = 0$ pointwise, we have $\lim_{n\to\infty}\langle u_n, \ell\rangle = 0$. This insures the σ-additivity of ℓ. Note that as $\operatorname{supp}\ell \subset \mathcal{S}$, for all $u \in C_c$ one can shortly write $\langle u, \ell\rangle$ instead of the meaningful bracket $\langle u_{|\mathcal{S}}, \ell\rangle$.

Unfortunately, this pointwise convergence of (u_n) is weaker than the uniform convergence with respect to which any $\ell \in C_c'$ is continuous. Except if $A\times B$ is compact, since in this special case, any décroissant sequence of continuous functions which converges pointwise to zero also converges uniformly on the compact space $\mathcal{S}$.

So far, we have only used the fact that $A\times B$ is a metric space. We now rely on the Polishness of A and B to get rid of this compactness restriction. It is known that any probability measure P on a Polish space X is tight (i.e. a Radon measure): for all $\epsilon > 0$, there exists a compact set $K_\epsilon \subset X$ such that $P(X \setminus K_\epsilon) \leq \epsilon$ ([5], Proposition II.7.3). As in addition a Polish space is completely regular, there exists a continuous function f_ϵ with a compact support such that $0 \leq f_\epsilon \leq 1$ and $\int_X (1-f_\epsilon)\,dP \leq \epsilon$. This is true in particular for the probability measures $\mu \in \mathcal{P}_A$ and $\nu \in \mathcal{P}_B$ which specify the constraint in (MK). Hence, there exist $\varphi_\epsilon \in C_A$ and $\psi_\epsilon \in C_B$ with compact supports such that $0 \leq \varphi_\epsilon, \psi_\epsilon \leq 1$ and $0 \leq \int_A (1-\varphi_\epsilon)\,d\mu, \int_B (1-\psi_\epsilon)\,d\nu \leq \epsilon$. It follows that any $\ell \in C_c'$ with $\ell_A = \mu$ and $\ell_B = \nu$ satisfies $0 \leq \langle (1-\varphi_\epsilon\otimes\psi_\epsilon), \ell\rangle \leq 2\epsilon$. With the following easy estimate $0 \leq \langle u_n, \ell\rangle \leq 2\epsilon + \langle u_n(\varphi_\epsilon\otimes\psi_\epsilon), \ell\rangle$ and the compactness of the support of $\varphi_\epsilon\otimes\psi_\epsilon$, one concludes that $\lim_{n\to\infty}\langle u_n, \ell\rangle = 0$ which proves (3.13).

• Proof of (d). As c is bounded below and lower semicontinuous on a metric space, it is the pointwise limit of a croissant sequence (c_n) of continuous bounded functions. It follows from the monotone convergence theorem that for any $\ell \in \mathcal{P}_{\mathcal{S}}$, $\Phi_1^*(\ell) = \int_{\mathcal{S}} c\,d\ell$. This completes the proof of the proposition. □

Optimal plan: an overview of the proofs of Theorems 3.3 and 3.5. The proofs of these theorems are postponed to Section 3.6. We first derive preliminary results at Sections 3.3, 3.4 and 3.5.

At Section 3.3, the abstract results of Section 2 are translated in terms of the Monge-Kantorovich problem. This is summarized at Theorem 3.24 which states an abstract characterization of the optimal plans. This theorem directly results from the extended saddle-point method. In particular, the optimal plan π is related to some linear form $\omega \in \mathcal{X}_1^*$. It remains to show that ω is the extension of some couple of functions (φ, ψ).

This is done at Section 3.4 for the sufficient condition and at Section 3.5 for the necessary condition. The main results of Sections 3.4 and 3.5 are respectively Lemma 3.31 and Lemma 3.43.

3.3. **Optimal plan: applying the extended saddle-point method.** The main result of this section is Theorem 3.24 which gives an abstract characterization of an optimal plan.

The space $\mathcal{X}_1$. By (2.3), we see that $|(\varphi,\psi)|_\Lambda = \|\varphi\oplus\psi\|_c$. This leads to

$$\mathcal{X}_1 = \{(\kappa_1,\kappa_2)\in C_A^*\times C_B^*; |(\kappa_1,\kappa_2)|_\Lambda^* < \infty\}$$

where $|(\kappa_1,\kappa_2)|_\Lambda^* = \sup\{\langle\varphi,\kappa_1\rangle + \langle\psi,\kappa_2\rangle; (\varphi,\psi)\in\mathcal{Y}_1, \|\varphi\oplus\psi\|_c\le 1\}$. The dual equality (2.8) gives

$$|(\kappa_1,\kappa_2)|_\Lambda^* = \inf\left\{\|\ell\|_c^*; \ell\in C_c' : \ell_A=\kappa_1, \ell_B=\kappa_2\right\}.$$

Note that $\mathcal{X}_1$ is the space of all $(\kappa_1,\kappa_2)\in C_A^*\times C_B^*$ such that $\kappa_1=\ell_A$ and $\kappa_2=\ell_B$ for some ℓ in C_c'. Recall that the elements of $\mathcal{X}_1=\mathcal{Y}_1'$ are identified with their restriction to $\mathcal{Y}$ which is dense in $\mathcal{Y}_1$.

Remark 3.14 (The space $\mathcal{Y}_1$ and the problem (D_1)). The exact description of $\mathcal{Y}_1$ and (D_1) will not be used later. Nevertheless, as an illustration of our general results, we describe them assuming that c is finitely valued. As $\Lambda(\varphi,\psi)=\delta_{\{\varphi\oplus\psi\le c\}}$, one sees that

$$\mathcal{Y}_1 = \{(\varphi,\psi); \varphi: A\to\mathbb{R} \text{ continuous}, \psi: B\to\mathbb{R} \text{ continuous} : \varphi\oplus\psi\in C_c\}$$

This result is not as obvious as it seems to be. It follows from an interesting paper [2] of J.M. Borwein and A.S. Lewis which studies the convergence of sequences of the form $(\varphi_n\oplus\psi_n)_{n\ge1}$. The additive form $\varphi\oplus\psi$ in the expression of $\mathcal{Y}_1$ is proved at ([2], Corollary 3.5) and the continuity of φ and ψ is a consequence of ([2], Proposition 5.1).

The corresponding problem (D_1) is

$$\begin{aligned}&\textit{maximize } \int_A \varphi\,d\mu + \int_B \psi\,d\nu \textit{ for all } \varphi,\psi \textit{ such that}\\ &\varphi,\psi \textit{ continuous, } \varphi\oplus\psi\in C_c \textit{ and } \varphi\oplus\psi\le c.\end{aligned}$$

Anyway, we won't use this dual problem since it is sandwiched between (D_0) and (D_2).

The extension Φ_2. To proceed, one has to compute the extension Φ_2. As it is the greatest convex $\sigma(C_c'^*, C_c')$-lower semicontinuous extension of Φ, we have

$$\Phi_2(\xi) = \delta_{\overline{\Gamma}}(\xi), \quad \xi\in C_c'^* \tag{3.15}$$

where $\overline{\Gamma}$ is the $\sigma(C_c'^*, C_c')$-closure of

$$\Gamma = \{u\le c\}\subset C_c.$$

Any $\omega\in\mathcal{X}_1^*$ is decomposed as $\omega=(\omega_A,\omega_B)$ where for all $(\kappa_1,\kappa_2)\in\mathcal{X}_1$, $\langle\omega,(\kappa_1,\kappa_2)\rangle = \langle\bar\omega,(\kappa_1,0)\rangle + \langle\bar\omega,(0,\kappa_2)\rangle = \langle\omega_A,\kappa_1\rangle + \langle\omega_B,\kappa_2\rangle$ where $\omega\in\mathcal{X}_1^*$ is seen as the restriction to $\mathcal{X}_1$ of some linear form $\bar\omega$ on $\mathcal{X} = C_A^*\times C_B^*$. The adjoint operator T_2^* is defined for all $\omega\in\mathcal{X}_1^*$ and $\ell\in C_c'$ by $\langle T_2^*\omega,\ell\rangle = \langle\omega_A,\ell_A\rangle + \langle\omega_B,\ell_B\rangle := \langle\omega_A\oplus\omega_B,\ell\rangle$. That is

$$T_2^*\omega = \omega_A\oplus\omega_B\in C_c'^*. \tag{3.16}$$

This yields

$$\Lambda_2(\omega) = \delta_{\overline{\Gamma}}(\omega_A\oplus\omega_B), \quad \omega\in\mathcal{X}_1^*$$

and the extended dual problem (D_2) is

$$\textit{maximize } \langle\omega_A,\mu\rangle + \langle\omega_B,\nu\rangle, \quad \omega\in\mathcal{X}_1^* \textit{ such that } \omega_A\oplus\omega_B\in\overline{\Gamma} \tag{D_2}$$

Note that for this dual problem to be meaningful, it is necessary that (H_C) holds: i.e. $(\mu,\nu)\in\mathcal{X}_1$. This is realized if $(\mu,\nu)\in\operatorname{dom}\Lambda^*$ or equivalently if $\inf(MK)<\infty$.

The constraint qualification. One will be allowed to apply Theorem 2.13 under the constraint qualification (2.14):

$$(\mu,\nu)\in\operatorname{diffdom}\Lambda^*. \tag{3.17}$$

Let us give some details on this abstract requirement.

Remark 3.18. Note that for all $\mu\in\mathcal{P}_A$, $\nu\in\mathcal{P}_B$, $(\mu,\nu)\notin\operatorname{icordom}\Lambda^*$ if $A\times B$ is an infinite set. Indeed, for all $\pi\in P(\mu,\nu)$ such that $\int_{A\times B}c\,d\pi<\infty$, one can find (a_o,b_o) such that with $\varepsilon_{(a_o,b_o)}$ the Dirac measure at (a_o,b_o), $\ell_t:=t\varepsilon_{(a_o,b_o)}+(1-t)\pi\not\geq 0$ for all $t<0$, so that $\Phi_1^*(\ell_t)=+\infty$ (Proposition 3.12-a). This shows that $[\ell_0,\ell_1]=[\pi,\varepsilon_{(a_o,b_o)}]\subset\operatorname{dom}\Phi_1^*$ while $\ell_t\notin\operatorname{dom}\Phi_1^*$ for all $t<0$. Hence, $(\mu,\nu)\notin\operatorname{icordom}\Lambda^*$ and one has to consider the assumption (3.17) on (μ,ν) rather than $(\mu,\nu)\in\operatorname{icordom}\Lambda^*$.
This is in contrast with the situation encountered in [4] where the rule is $x_o\in\operatorname{icordom}\Lambda^*$.

Lemma 3.19. *We have* $\operatorname{dom}\Lambda^*=\operatorname{diffdom}\Lambda^*$.

Proof. Proposition 3.12-a states that $\operatorname{dom}\Phi_1^*\subset\mathcal{L}^+$ where $\mathcal{L}^+=\{\ell\in C_c';\ell\geq 0\}$ is the the positif cone of C_c'. Therefore, $\Phi_1^*=\Phi_1^*+\delta_{\mathcal{L}^+}$. Consequently, with (2.7) one obtains that

$$\Lambda^*(x)=\inf\{\Phi_1^*(\ell);\ell\in\mathcal{L}^+,T\ell=x\},\quad x\in\mathcal{X}_1.$$

Suppose ad absurdum that there is some $x_o\in\operatorname{dom}\Lambda^*$ such that $x_o\notin\operatorname{diffdom}\Lambda^*$. This implies that there exists some half-line $]x_o,x_o+\infty(x_o'-x_o)[$ on which Λ^* achieves the value $+\infty$, which in turn implies that Φ_1^* must achieve the value $+\infty$ somewhere on $\mathcal{L}^+$. But this is impossible since $\Phi_1^*(\ell)=\|\ell\|_c^*$ for all $\ell\in\mathcal{L}^+$. This completes the proof of the lemma. ☐

As a consequence of this lemma, it appears that (3.17) is *not* a constraint qualification. One can apply Theorem 2.13 under the only restriction that $\inf(MK)<\infty$. This gives the following

Lemma 3.20. *Let us assume that* $\inf(MK)<\infty$. *Then,* (P) *and* (D_2) *both admit a solution in* $\mathcal{P}_{AB}$ *and* $\mathcal{X}_1^*$. *Furthermore, any* $(\pi,\omega)\in\mathcal{P}_{AB}\times\mathcal{X}_1^*$ *is a solution of* (P) *and* (D_2) *if and only if*

$$\begin{cases}(a)\quad \pi_A=\mu,\pi_B=\nu,\int_{A\times B}c\,d\pi<\infty;\\ (b)\quad \pi\in\partial_{C_c'}\Phi_2(\eta)\ \textit{where}\\ (c)\quad \eta=T_2^*\omega.\end{cases} \tag{3.21}$$

As Φ_1^* and Φ_2 are mutually convex conjugates, (3.21-b) is equivalent to

$$\eta\in\partial_{C_c'^*}\Phi_1^*(\pi) \tag{3.22}$$

and also equivalent to Young's identity

$$\Phi_1^*(\pi)+\Phi_2(\eta)=\langle\eta,\pi\rangle \tag{3.23}$$

and also equivalent to

$$\begin{cases}\Phi_2(\eta)=0\\ \langle\eta,\pi\rangle=\int_{A\times B}c\,d\pi.\end{cases}$$

In other words:

Theorem 3.24. *Let* $\pi\in P(\mu,\nu)$ *be such that* $\int_{A\times B}c\,d\pi<\infty$. *Then:*

(1) (D_2) *admits at least a solution in* $\mathcal{X}_1^*$;

(2) π is an optimal plan if and only if there exists some $\omega \in \mathcal{X}_1^*$ such that

$$\begin{cases} (a) & T_2^*\omega \in \overline{\Gamma} \\ (b) & \langle \omega, (\mu, \nu) \rangle = \int_{A \times B} c\, d\pi; \end{cases}$$

With $\eta = T_2^*\omega$, this implies the equivalent statements (3.21-b), (3.22) or (3.23).

(3) If such an ω exists, it is a solution of (D_2) and any other solution of (D_2) is also convenient.

This is the core of the extended saddle-point method applied to Monge-Kantorovich problem. To prove a practical optimality criterion one still has to translate these abstract properties.

3.4. **Optimal plan: preliminary results for the sufficient condition.** The next lemmas are preliminary results for the proof of a sufficient condition for the optimality.

Lemma 3.25. *Let φ and ψ be real functions on A and B.*

(1) The lower semicontinuous regularizations $\mathrm{ls}\,\varphi$ and $\mathrm{ls}\,\psi$ of φ and ψ satisfy

$$\mathrm{ls}\,(\varphi \oplus \psi) = \mathrm{ls}\,\varphi \oplus \mathrm{ls}\,\psi.$$

(2) If φ and ψ are such that $\varphi \oplus \psi = c$ on some subset $\mathcal{T}$ of $A \times B$ and $\varphi \oplus \psi \le c$ everywhere on $A \times B$. Then, $\mathrm{ls}\,\varphi$ and $\mathrm{ls}\,\psi$ still share the same properties.

Proof. • Proof of (1). For each $(a, b) \in A \times B$,

$$\begin{aligned} \mathrm{ls}\,(\varphi \oplus \psi)(a, b) &= \sup_{V \in \mathcal{N}((a,b))} \inf_{(a',b') \in V} [\varphi(a') + \psi(b')] \\ &= \sup \left\{ \inf_{(a',b') \in V_A \times V_B} [\varphi(a') + \psi(b')]; V_A \in \mathcal{N}(a), V_B \in \mathcal{N}(b) \right\} \\ &= \sup \left\{ \inf_{a' \in V_A} \varphi(a'); V_A \in \mathcal{N}(a) \right\} + \sup \left\{ \inf_{b' \in V_B} \psi(b'); V_B \in \mathcal{N}(b) \right\} \\ &= \mathrm{ls}\,\varphi(a) + \mathrm{ls}\,\psi(b) \end{aligned}$$

where $\mathcal{N}(x)$ stands for the set of all open neighbourhoods of x.

• Proof of (2). It is a direct consequence of the lower semicontinuity of c and statement (1). □

Lemma 3.26. *Let $\pi \in P(\mu, \nu)$ be such that $\int_{A \times B} c\, d\pi < \infty$. Suppose that there exists two real-valued functions $\varphi \in \mathbb{R}^A$ and $\psi \in \mathbb{R}^B$ such that*

$$\begin{cases} \varphi \oplus \psi \le c & \text{everywhere} \\ \varphi \oplus \psi = c & \pi\text{-almost everywhere.} \end{cases} \tag{3.27}$$

(1) If φ is μ-measurable and ψ is ν-measurable, then $\varphi \in L_1(A, \mu)$ and $\psi \in L_1(B, \nu)$.

(2) In any case, the real-valued functions $\mathrm{ls}\,\varphi$ and $\mathrm{ls}\,\psi$ still satisfy (3.27) together with $\mathrm{ls}\,\varphi \in L_1(A, \mu)$ and $\mathrm{ls}\,\psi \in L_1(B, \nu)$.

Proof. • Proof of (1). Let us fix $(a_o, b_o) \in A \times B$ such that $c(a_o, b_o) < \infty$ (such a point exists since $\int_{A \times B} c\, d\pi < \infty$ for some π.) We have $\varphi(a) = c(a, b_o) - \psi(b_o) \ge -\psi(b_o)$ for all $a \in A$ and similarly $\psi \ge -\varphi(a_o)$. Hence, the integrals $\int_A \varphi\, d\mu \in [-\psi(b_o), +\infty]$ and $\int_B \psi\, d\nu \in [-\varphi(a_o), +\infty]$ are well-defined. Finally, $\varphi \in L_1(A, \mu)$ and $\psi \in L_1(B, \nu)$ since $\int_A \varphi\, d\mu + \int_B \psi\, d\nu = \int_{A \times B} c\, d\pi < \infty$.

• Proof of (2). Applying Lemma 3.25 with $\mathcal{T}$ a measurable set such that $\pi(\mathcal{T}) = 1$ yields two lower bounded measurable functions $\mathrm{ls}\,\varphi$ and $\mathrm{ls}\,\psi$ which still satisfy (3.27). One concludes as above. □

Let $\overline{\Upsilon}$ be the $\sigma(\mathcal{X}_1^*, \mathcal{X}_1)$-closure of

$$\Upsilon = \{(\varphi, \psi) \in C_A \times C_B; \varphi \oplus \psi \leq c\}. \tag{3.28}$$

Lemma 3.29.
(a) For all $(a, b) \in \mathcal{S}$, $\Lambda^*(\varepsilon_a, \varepsilon_b) = c(a, b)$.
(b) For any $\omega \in \mathcal{X}_1^*$, *we have* $\omega \in \overline{\Upsilon}$ *if and only if* $\langle \omega, \kappa \rangle \leq \Lambda^*(\kappa), \forall \kappa \in \mathcal{X}_1$.
(c) $T_2^* \overline{\Upsilon} \subset \overline{\Gamma}$.

Proof. • Proof of (a). For any $(a, b) \in \mathcal{S}$, $\Lambda^*(\varepsilon_a, \varepsilon_b) = \inf\{\int_{A\times B} c\, d\pi; \pi \in \mathcal{P}_{AB} : \pi_A = \varepsilon_a, \pi_B = \varepsilon_b\} = \int_{A\times B} c\, d\varepsilon_{(a,b)} = c(a, b)$ where we used the dual equality (2.7) and the fact that $\varepsilon_{(a,b)}$ is the unique plan π with marginals ε_a and ε_b.

• Proof of (b). It is enough to check that for all $\phi = (\varphi, \psi)$ in $C_A \times C_B$

$$\phi \in \Upsilon \Leftrightarrow [\langle \phi, \kappa \rangle \leq \Lambda^*(\kappa), \forall \kappa \in \mathcal{X}_1]. \tag{3.30}$$

Young's inequality $\langle \phi, \kappa \rangle \leq \Lambda(\phi) + \Lambda^*(\kappa), \forall \phi, \kappa$ and $\phi \in \Upsilon \Leftrightarrow \Lambda(\phi) = \delta_\Upsilon(\phi) = 0$ give the direct implication. For the converse, choosing $\kappa = (\varepsilon_a, \varepsilon_b)$ in the right-hand side of (3.30), one obtains with the previous statement (a) that $\varphi \oplus \psi \leq c$.

• Proof of (c). It is clear that $T^*\Upsilon \subset \Gamma$ and one concludes with the $\sigma(\mathcal{X}_1^*, \mathcal{X}_1)$-$\sigma(\mathcal{L}_1^*, \mathcal{L}_1)$-continuity of $T_2^* : \mathcal{X}_1^* \to \mathcal{L}_1^*$, see Lemma 4.1-d. □

Lemma 3.31. *Let* $\pi \in P(\mu, \nu)$ *be such that* $\int_{A\times B} c\, d\pi < \infty$ *and suppose that there exist two real functions* φ *in* $L_1(A, \mu)$ *and* ψ *in* $L_1(B, \nu)$ *satisfying (3.27).*
Then, there exists some ω *in* $\overline{\Upsilon}$ *such that*

$$\begin{cases} \langle \omega, (\mu, \nu) \rangle = \int_{A\times B} c\, d\pi; \\ \omega(\epsilon_a, \epsilon_b) = \varphi(a) + \psi(b) \textit{ for } \pi\textit{-a.e. } (a, b) \textit{ and} \\ \omega(\kappa) \leq \Lambda^*(|\kappa|), \ \forall \kappa \in \mathcal{X}_1. \end{cases}$$

Proof. There exists a measurable subset $\mathcal{T}$ of $\mathcal{S}$ such that $\pi(\mathcal{T}) = 1$ and $\varphi \oplus \psi = c$ everywhere on $\mathcal{T}$. Let E_o be the vector subspace of $\mathcal{X}_1$ spanned by (μ, ν) and $\{(\epsilon_a, \epsilon_b); (a, b) \in \mathcal{T}\}$. It follows from our assumptions on φ and ψ that for all positif $\kappa = (\kappa_1, \kappa_2) \in E_o$, φ is in $L_1(A, \kappa_1)$ and ψ is in $L_1(B, \kappa_2)$. Define the linear form ω_o on E_o for each $\kappa \in E_o$ by

$$\omega_o(\kappa) = \int_A \varphi\, d\kappa_1 + \int_B \psi\, d\kappa_2.$$

Clearly,

$$\omega_o(\mu, \nu) = \int_{A\times B} c\, d\pi \tag{3.32}$$

and for all positif $\kappa \in E_o$,

$$\begin{aligned} \omega_o(\kappa) &= \int_A \varphi\, d\kappa_1 + \int_B \psi\, d\kappa_2 \\ &\leq \sup\left\{ \int_A \tilde{\varphi}\, d\kappa_1 + \int_B \tilde{\psi}\, d\kappa_2;\ \tilde{\varphi} \in L_1(A, \kappa_1),\ \tilde{\psi} \in L_1(B, \kappa_1),\ \tilde{\varphi} \oplus \tilde{\psi} \leq c \right\}. \end{aligned}$$

Denoting (K_κ) and $(\overline{K}_\kappa)$ the analogues of problems (K) and $(\overline{K})$ with (κ_1, κ_2) instead of (μ, ν), this means that

$$\omega_o(\kappa) \leq \sup(\overline{K}_\kappa).$$

The dual equality (2.8) states that $\sup(K_\kappa) = \Lambda^*(\kappa)$. As we have already seen at Theorem 3.2-a that $\sup(\overline{K}_\kappa) = \sup(K_\kappa)$, we obtain: $\sup(\overline{K}_\kappa) = \Lambda^*(\kappa)$. Therefore, we have proved that $\omega_o(\kappa) \le \Lambda^*(\kappa)$, for all $\kappa \in E_o$, $\kappa \ge 0$. As for any $\kappa \in E_o$, $\omega_o(\kappa) = \int_{A\times B} c\, d\rho$ for any measure ρ with marginals $\rho_A = \kappa_1$ and $\rho_B = \kappa_2$, one sees that ω_o is positif. It follows that

$$\omega_o(\kappa) \le \Lambda^*(|\kappa|), \quad \kappa \in E_o$$

where $|\kappa| = (|\kappa_1|, |\kappa_2|)$ and $|\kappa_i|$ is the absolute value of the measure κ_i.
Note that $\mathcal{X}_1$ is a Riesz space since it is the topological dual of a normed Riesz space. Hence, any $\kappa \in \mathcal{X}_1$ admits positif and négatif parts κ^+ and κ^-, and its absolute value is $|\kappa| = \kappa^+ + \kappa^-$. This allows to consider the positively homogeneous convex function $\Lambda^*(|\kappa|)$ on the vector space E_1 spanned by $\operatorname{dom}\Lambda^*$. By the analytic form of Hahn-Banach theorem, there exists an extension ω of ω_o to E_1 which satisfies $\omega(\kappa) \le \Lambda^*(|\kappa|)$ for all $\kappa \in E_1$. But $E_1 = \mathcal{X}_1$ and one completes the proof of the lemma with (3.32) and Lemma 3.29-b. $\square$

3.5. **Optimal plan: preliminary results for the necessary condition.** Under the condition (3.21-a), π necessarily satisfies: $\operatorname{supp}\pi \subset \mathcal{S}$. This fact will be invoked without warning.

Lemma 3.33. *Let π and η satisfy (3.21-a,b). Then, the restriction of η to $L_\infty.\pi := \{h.\pi; h \in L_\infty(A\times B, \pi)\}$ is given by*

$$\langle \eta, h.\pi\rangle = \int_{A\times B} hc\, d\pi, \quad \forall h \in L_\infty(\pi) \tag{3.34}$$

Proof. To specify the restriction γ of η to $L_\infty.\pi$, it is enough to vary Φ_1^* in the direction $L_\infty.\pi$ to get with (3.22): $\gamma \in \partial_{(L_\infty.\pi)^*}\Phi_1^*(\pi)$. Taking $h \in L_\infty(\pi)$ such that $\|h\|_\infty \le 1$, by monotone convergence we obtain $\Phi_1^*(\pi + h.\pi) = \sup\{\int_{A\times B}(1+h)u\, d\pi; u \in C_c, u \le c\} = \int_{\mathcal{S}}(1+h)c\, d\pi$. It comes out that $\partial_{(L_\infty.\pi)^*}\Phi_1^*(\pi) = \{c\}$, which gives (3.34). $\square$

We first derive the necessary condition in the special case where c is assumed to be finite and continuous.

Proposition 3.35. *Assume that c is finite and continuous and let π be an optimal plan. Then, there exist two finitely-valued upper semicontinuous functions φ on A and ψ on B such that*

$$\begin{cases} \varphi \oplus \psi \le c & \textit{everywhere and} \\ \varphi \oplus \psi = c & \textit{on } \operatorname{supp}\pi. \end{cases}$$

Proof. At the beginning of this proof, c is only assumed to be finite and lower semicontinuous. By Lemma 3.19, (3.17) is satisfied. Let η and ω be as in (3.21-b & c). Because of Theorem 2.13-c-1 & 2, there exists a generalized sequence $\{(\alpha_\tau, \beta_\tau)\}$ in $\operatorname{dom}\Lambda_1$ such that $\lim_\tau T_1^*(\alpha_\tau, \beta_\tau) = \eta$ with respect to $\sigma(C_c'^*, C_c')$. As $T_1^*\operatorname{dom}\Lambda_1 \subset \mathcal{U}_1$ (see Lemma 4.1-g), $T_1^*(\alpha_\tau, \beta_\tau) = \alpha_\tau \oplus \beta_\tau \in C_c$ and

$$\begin{cases} (a) & \lim_\tau \alpha_\tau \oplus \beta_\tau = \eta \text{ with} \\ (b) & C_c \ni \alpha_\tau \oplus \beta_\tau \le c \text{ for all } \tau \end{cases} \tag{3.36}$$

Defining

$$\widetilde{\eta}(a,b) = \langle \eta, \epsilon_{(a,b)}\rangle, \quad (a,b) \in A\times B, \tag{3.37}$$

where $\epsilon_{(a,b)}$ is the Dirac mass at (a,b), one immediately sees that

$$\widetilde{\eta} \le c. \tag{3.38}$$

Furthermore, since $\widetilde{\eta} = \lim_\tau \alpha_\tau \oplus \beta_\tau$ pointwise (C'_c contains the Dirac masses), by ([2], Corollary 3.5) we obtain that $\widetilde{\eta} = \varphi \oplus \psi$ for some functions φ and ψ on A and B. This gives us some hope to complete the proof, but as will be seen below, $\widetilde{\eta}$ isn't the right function to be considered.
For any (a, b) in $\operatorname{supp} \pi$: the support of π, one can find a sequence $\{h_k\}$ in C_c such that $\lim_k h_k.\pi = \epsilon_{(a,b)}$ in $\mathcal{P}_c$, see Lemma 3.60 below. As c is lower semicontinuous, with (3.34) we obtain

$$\liminf_k \langle h_k.\pi, \eta\rangle = \liminf_k \int_{A\times B} h_k c\, d\pi \geq c(a, b). \tag{3.39}$$

Unfortunately, no regularity property for η has been established to insure that $\langle \eta, \epsilon_{(a,b)}\rangle \geq \liminf_k \langle \eta, h_k.\pi\rangle$; this would lead to the converse of (3.38): $\widetilde{\eta} \geq c$ on $\operatorname{supp} \pi$. An alternate strategy is to introduce the upper semicontinuous regularization

$$\overline{\eta} = \operatorname{us} \widetilde{\eta}$$

of $\widetilde{\eta}$ on $A\times B$. As $\overline{\eta}$ is upper semicontinuous, for all $(a, b) \in \operatorname{supp} \pi$, we have $\overline{\eta}(a, b) \geq \limsup_k \int_{A\times B} \overline{\eta} h_k\, d\pi$. Now, one obtains with (3.39) that

$$\overline{\eta}(a, b) \geq c(a, b), \quad \forall (a, b) \in \operatorname{supp} \pi. \tag{3.40}$$

Regularizing both sides of (3.38) and assuming that c is upper semicontinuous and therefore *continuous*, we obtain that

$$\overline{\eta} \leq c. \tag{3.41}$$

It remains to check that

$$\overline{\eta} = \varphi \oplus \psi$$

for some finitely-valued upper semicontinuous functions φ and ψ on A and B. With (3.16) and (3.21) we know that $\eta = \omega_A \oplus \omega_B$ for some $\omega \in \mathcal{X}_1^*$. It follows that $\widetilde{\eta} = \widetilde{\omega}_A \oplus \widetilde{\omega}_B$ where $\widetilde{\omega}_A(a) = \omega_A(\epsilon_a)$ and $\widetilde{\omega}_B(b) = \omega_B(\epsilon_b)$. With Lemma 3.25, one sees that $\overline{\eta} = \operatorname{us} \widetilde{\omega}_A \oplus \operatorname{us} \widetilde{\omega}_B$. This proves the desired result with $\varphi = \operatorname{us} \widetilde{\omega}_A$ and $\psi = \operatorname{us} \widetilde{\omega}_B$. Since $\widetilde{\eta} \leq \overline{\eta} \leq c$ and both $\widetilde{\eta}$ and c are finitely-valued, so are φ and ψ. □

Remark 3.42. By means of the usual approaches [8, 1, 10], one can prove when c is finitely-valued that under the assumptions (1.6) or (1.7), φ and ψ can be required to be c-concave conjugates to each other. In the special case where c is assumed to be continuous, c-concave conjugates are upper semicontinuous. This is in accordance with Proposition 3.35.

Now, we remove the assumption that c is finite and continuous and only assume that it is lower semicontinuous. The main technical result for the proof of the characterization of the optimal plans is the following

Lemma 3.43. *Assume that c is a $[1, \infty]$-valued lower semicontinuous function. Let $\pi \in \mathcal{P}_c$ and $\eta \in C'^*_c$ be as in (3.21-b), i.e. $\pi \in \partial_{C'_c} \Phi_2(\eta)$ and define the function $\widetilde{\eta}$ on $\mathcal{S}$ by*

$$\widetilde{\eta}(a, b) = \langle \eta, \epsilon_{(a,b)}\rangle, \quad (a, b) \in \mathcal{S}. \tag{3.44}$$

Then,

$$\begin{cases} \widetilde{\eta} \leq c & \text{on } \mathcal{S} \\ \widetilde{\eta} = c & \text{on } \operatorname{supp} \pi \bigcap \{c < \infty\} \end{cases}$$

and $\widetilde{\eta}$ is a finitely-valued measurable function on $\mathcal{S}$.

Remark 3.45. We assume that $c \geq 1$ without loss of generality, see Remark 3.8, to allow dividing by c in the definition of C_c.

Proof. Because of Theorem 2.13-c-2, there exists a sequence $\{\rho_n\}$ in C_{AB} such that

$$\begin{cases} \rho_n \le c, \forall n \quad \text{and} \\ \lim_{n\to\infty} \rho_n = \eta \end{cases} \tag{3.46}$$

with respect to $\sigma(C_c'^*, C_c')$. Having Remark 3.9 in mind, recall that only the restriction of ρ_n to $\mathcal{S}$ carries information as regards to the dual pairing $\langle C_c'^*, C_c' \rangle$. Also recall that the k^{th} *Moreau-Yosida approximation* of a function u on a space with metric d is defined for all x by $u^{(k)}(x) = \inf_y \{u(y) + kd(x,y)\}$. Defining the Moreau-Yosida approximations

$$\begin{aligned} \rho_{n,k} &= [\max(\rho_n, k)]^{(k)} \\ c_k &= [\max(c,k)]^{(k)} \end{aligned}$$

for all $n, k \ge 1$, (3.46) implies that

$$\begin{cases} (a) & \rho_{n,k}, c_k \in C_{AB},\ \rho_{n,k} \le c_k \le c,\ \forall n \ge 1 \\ (b) & \lim_n \lim_k \langle \rho_{n,k}, m \rangle = \langle \eta, m \rangle,\ \forall m \in \mathcal{M}_c, m \ge 0 \\ (c) & 0 \le c_k \uparrow c \text{ pointwise.} \end{cases} \tag{3.47}$$

where $\mathcal{M}_c$ is the space of all measures m on $\mathcal{S}$ such that $\int_{\mathcal{S}} c\, d|m| < \infty$. By Remark 3.9, one sees that $\mathcal{M}_c \subset C_c'$.

While deriving (3.47), we used the well-known results:

- a Moreau-Yosida approximation is a continuous function and
- the sequence of Moreau-Yosida approximations of a function tends pointwise and croissantly to its lower semicontinuous regularization.

The proof of statement (3.47-b) relies on the monotone convergence theorem; this is the reason why it holds for all m in $\mathcal{M}_c$ rather than in C_c'.

Let us introduce the cone $\mathcal{Q}^+ = \{\ell \in C_c'; \ell \ge 0 \text{ and } \langle \eta, \ell \rangle \ge 0\}$ and $\mathcal{Q}$ the vector space spanned by $\mathcal{Q}^+$. We first consider the restriction θ of η to $\mathcal{Q}$. By (3.34), π is in $\mathcal{Q}^+$ and (3.22) gives us $\theta \in \partial_{\mathcal{Q}^*} N(\pi)$ where $N(\ell) = \sup\{\langle u, \ell \rangle; u \in C_c, |u| \le c\}$, $\ell \in \mathcal{Q}$ which is the dual norm $\|\cdot\|_c^*$ restricted to $\mathcal{Q}$. It follows that θ belongs to the topological dual space $\mathcal{Q}'$ of the normed space $(\mathcal{Q}, \|\cdot\|_c^*)$:

$$\theta := \eta_{|\mathcal{Q}} \in \mathcal{Q}'.$$

This topological regularity of θ will allow us a few lines below to invoke Brønsted-Rockafellar's lemma. It is not clear that η is continuous on the whole normed space C_c'.

Let us denote Ψ the restriction of Φ_1^* to $\mathcal{Q}$ and Ψ^* its convex conjugate with respect to the dual pairing $\langle \mathcal{Q}, \mathcal{Q}' \rangle$. Since $\theta \in \mathcal{Q}'$, $\theta \ge 0$ and $\eta \in \overline{\Gamma}$, one sees that $0 \le \Psi^*(\theta) \le \Phi_2(\eta) = 0$. As (3.21-b) is equivalent to Young's identity (3.23), one obtains

$$\Psi(\pi) + \Psi^*(\theta) = \langle \theta, \pi \rangle = \lim_{n\to\infty} \langle \xi_n, \pi \rangle$$

where

$$\xi_n = \rho_{n,k(n)|\mathcal{Q}} \in \mathcal{Q}' \tag{3.48}$$

is the restriction of $\rho_{n,k(n)} \in C_c'^*$ to $\mathcal{Q} \subset C_c'$ for some sequence $\{k(n)\}_n$ which converges fast enough to infinity to imply that $\lim_{n\to\infty} \langle \xi_n, \pi \rangle = \langle \theta, \pi \rangle$ by means of (3.47-b).

Denote Ψ_n the restriction to $\mathcal{Q}$ of the analogue of Φ_1^* with $c_{k(n)}$ instead of c and Ψ_n^* its convex conjugate with respect to $\langle \mathcal{Q}, \mathcal{Q}' \rangle$. By (3.47-c), we have $\lim_{n\to\infty} \Psi_n(\pi) = \Psi(\pi)$. By (3.47-a), we also have $\Psi^*(\theta) = \Psi_n^*(\xi_n) = 0$ for all n. Therefore,

$$\Psi_n(\pi) + \Psi_n^*(\xi_n) = \langle \xi_n, \pi \rangle + \epsilon_n$$

with $\lim_{n\to\infty}\epsilon_n = 0$. In other words, ξ_n is an ϵ_n-subgradient of Ψ_n at π. Hence, by Brønsted-Rockafellar lemma, there exist two sequences $\{\pi_n\}$ in $\mathcal{Q}$ and $\{\theta_n\}$ in $\mathcal{Q}'$ such that for all n,

$$\|\pi_n - \pi\| \le \sqrt{\epsilon_n}, \tag{3.49}$$

$$\|\xi_n - \theta_n\| \le \sqrt{\epsilon_n} \tag{3.50}$$

(both norms $N(\ell) = \sup\{\langle u, \ell\rangle; u \in C_c, |u| \le c\}$ on $\mathcal{Q}$ and $\sup\{\langle\cdot, \ell\rangle; \ell \in \mathcal{Q}, N(\ell) \le 1\}$ on $\mathcal{Q}'$ are simply written $\|\cdot\|$) and

$$\theta_n \in \partial_{\mathcal{Q}'}\Psi_n(\pi_n). \tag{3.51}$$

We define

$$G = \{(a,b) \in \mathcal{S}; \langle\eta, \epsilon_{(a,b)}\rangle \ge 0\}$$

the set of all $(a,b) \in A\times B$ such that $\epsilon_{(a,b)} \in \mathcal{Q}$. Since $c_{k(n)}$ is finite and continuous, proceeding as in Proposition 3.35, one shows as for (3.40) that

$$\overline{\theta}_n(a,b) = c_{k(n)}(a,b), \quad \forall (a,b) \in \mathrm{cl}\, G \cap \mathrm{supp}\, \pi_n \tag{3.52}$$

where

$$\overline{\theta}_n(a,b) = \mathrm{us}\, \widetilde{\theta}_n(a,b), \quad (a,b) \in \mathcal{S}$$

is the upper semicontinuous regularization of

$$\widetilde{\theta}_n(a,b) = \begin{cases} \theta_n(\epsilon_{(a,b)}) & \text{if } (a,b) \in G \\ -\infty & \text{otherwise} \end{cases}, \quad (a,b) \in \mathcal{S}$$

and $\mathrm{cl}\, G$ is the closure of G in $A\times B$. Since $\mathcal{S}$ is closed, we have $\mathrm{cl}\, G \subset \mathcal{S}$. As π_n may not be a measure, one uses Lemma 3.60 below instead of its usual analogue.

Thanks to (3.49), $\lim_{n\to\infty}\pi_n = \pi$ *strongly* in $\mathcal{Q}$ and for all large enough n we have $\pi_n \in \mathcal{Q}^+$ and

$$\mathrm{supp}\, \pi_n = \mathrm{supp}\, \pi. \tag{3.53}$$

With $\overline{\eta}$ the upper semicontinuous regularization of $\widetilde{\eta}$, we have $\{\overline{\eta} \ge 0\} = \mathrm{cl}\,\{\widetilde{\eta} \ge 0\}$. It follows from (3.40) that

$$\mathrm{supp}\, \pi \subset \mathrm{cl}\, G. \tag{3.54}$$

Thanks to (3.48) and (3.50), for all $0 \le r < \infty$,

$$\lim_{n\to\infty} \sup_{(a,b)\in G\cap\{c\le r\}} |\rho_{n,k(n)}(a,b) - \widetilde{\theta}_n(a,b)| = 0. \tag{3.55}$$

Let us assume for a while that

$$\sup_{\mathcal{S}} c < \infty. \tag{3.56}$$

Under this assumption, (3.55) leads us to $\lim_{n\to\infty}\sup_G |\rho_{n,k(n)} - \widetilde{\theta}_n| = 0$ on G. Upper regularizing, because of this uniform estimate and the continuity of $\rho_{n,k(n)}$, one obtains

$$\lim_{n\to\infty} \gamma_n = 0 \tag{3.57}$$

where $\gamma_n = \sup_{\mathrm{cl}\, G} |\rho_{n,k(n)} - \overline{\theta}_n|$. By (3.52), (3.53), (3.54) and $\rho_{n,k} \le \rho_n$ for all n, k, we obtain for all large enough $n : \rho_n(a,b) \ge c_{k(n)}(a,b) - \gamma_n$, $\forall (a,b) \in \mathrm{supp}\, \pi$. Letting n tend to infinity, we see with (3.46), (3.47-c) and (3.57) that $\widetilde{\eta}(a,b) = c(a,b)$, for all $(a,b) \in \mathrm{supp}\, \pi$ where $\widetilde{\eta}$ is defined at (3.44). We have just proved that under the assumption (3.56),

$$\begin{cases} (a) & \widetilde{\eta}(a,b) \le c(a,b), \quad \forall (a,b) \in \mathcal{S} \\ (b) & \widetilde{\eta}(a,b) = c(a,b), \quad \forall (a,b) \in \mathrm{supp}\, \pi \end{cases} \tag{3.58}$$

where the statement (a) directly follows from (3.46).

It remains to remove the restriction (3.56). For each $k \geq 1$, let

$$\left\{ \begin{array}{lcl} \mathcal{S}_k & = & \{c \leq k\} \quad \text{and} \\ c_k & = & c + \delta_{\mathcal{S}_k} \end{array} \right.$$

The function c_k is lower semicontinuous on $A \times B$ and satisfies (3.56); $\{\mathcal{S}_k\}$ is a croissant sequence of closed level sets of c with $\mathcal{S}_k \subset \mathcal{S}$ for all k. By Proposition 3.12-b we have $\operatorname{supp} \pi \subset \mathcal{S}$.
It is assumed that $\pi \in \partial_{C'_c} \Phi_2(\eta)$ which is equivalent to the Young's identity $\Phi_1^*(\pi) + \Phi_2(\eta) = \langle \eta, \pi \rangle$ or equivalently $[\Phi_1^*(\pi) = \langle \eta, \pi \rangle$ and $\Phi_2(\eta) = 0]$ which is also equivalent to

$$\left\{ \begin{array}{l} \langle \eta, \pi \rangle = \int_{A \times B} c \, d\pi \quad \text{and} \\ \langle \eta, \ell \rangle \leq \Phi_1^*(\ell), \ \forall \ell \in C'_c \end{array} \right. \tag{3.59}$$

because of Proposition 3.12-d and Lemma 3.61 below. Let us consider for each k

$$\left\{ \begin{array}{rcl} \pi_k & = & \mathbf{1}_{\mathcal{S}_k} . \pi \\ \langle \eta_k, \ell \rangle & = & \langle \eta, \ell \rangle, \quad \forall \ell \in C'_c \text{ such that } \operatorname{supp} \ell \subset \mathcal{S}_k \end{array} \right.$$

Note with Remark 3.9 that $\pi_k \in C'_{c_k}$ and $\eta_k \in C'^*_{c_k}$. Also introduce Θ_k and Θ_k^* the analogues of Φ_1^* and Φ_2 where c is replaced by c_k.
By Proposition 3.12-d and (3.34), $\Theta_k(\pi_k) = \int_{A \times B} c_k \, d\pi_k = \langle \eta_k, \pi_k \rangle$. Since for any $\ell \in C'_c$ such that $\operatorname{supp} \ell \subset \mathcal{S}_k$ we have $\Theta_k(\ell) = \Phi_1^*(\ell)$, one obtains with (3.59) that $\langle \eta_k, \ell \rangle \leq \Theta_k(\ell)$ for all $\ell \in C'_{c_k}$. Reasoning as for the derivation of (3.59) but taking the reverse way, this shows that

$$\pi_k \in \partial_{C'_{c_k}} \Theta_k^*(\eta_k).$$

Applying (3.58) yields

$$\left\{ \begin{array}{ll} \widetilde{\eta}_k(a,b) \leq c_k(a,b), & \forall (a,b) \in \mathcal{S}_k \\ \widetilde{\eta}_k(a,b) = c_k(a,b), & \forall (a,b) \in \operatorname{supp} \pi_k \end{array} \right. , \quad \forall k \geq 1$$

with $\widetilde{\eta}_k(a,b) = \langle \eta_k, \epsilon_{(a,b)} \rangle$, $(a,b) \in \mathcal{S}_k$. As $\cup_k \mathcal{S}_k = \{c < \infty\}$, this is equivalent to

$$\left\{ \begin{array}{ll} \widetilde{\eta}(a,b) \leq c(a,b), & \forall (a,b) \in \{c < \infty\} \\ \widetilde{\eta}(a,b) = c(a,b), & \forall (a,b) \in \operatorname{supp} \pi \bigcap \{c < \infty\} \end{array} \right.$$

Finally, one sees with (3.46) that $\widetilde{\eta} = \lim_{n \to \infty} \rho_n$ on $\mathcal{S}$. This implies that $\widetilde{\eta}$ is measurable on $\mathcal{S}$ and completes the proof of the lemma. □

During this proof, we have used the following elementary lemmas.

Lemma 3.60. *Let ℓ be a positif element of C'_c. For any $(a,b) \in \operatorname{supp} \ell$, there exists a sequence $\{h_k\}_{k \geq 1}$ of positif continuous bounded functions on $A \times B$ such that $\lim_{k \to \infty} h_k . \ell = \epsilon_{(a,b)}$ in with respect to $\sigma(C'_c, C_c)$.*

Proof. To see this, consider a décroissant sequence $\{G_k\}_{k \geq 1}$ of neighbourhoods of (a,b) with $\lim_k G_k = \{(a,b)\}$ and choose h_k such that $\{h_k > 0\} \subset G_k$ and $\langle h_k, \ell \rangle = 1$, this is possible since $A \times B$ is a metric space. □

Lemma 3.61. *For any $\eta \in C'^*_c$, the three following statements are equivalent:*

(i) $\Phi_2(\eta) = 0$;
(ii) $\eta \in \overline{\Gamma}$;
(iii) $\langle \eta, \ell \rangle \leq \Phi_1^*(\ell)$, *for all* $\ell \in C'_c$.

Proof. The equivalence (i) $\Leftrightarrow$ (ii) is an immediate consequence of (3.15).
Let us prove: (ii) $\Leftrightarrow$ (iii). Taking the closure, it is enough to check that for all u in C_{AB}

$$u \in \Gamma \Leftrightarrow [\langle u, \ell\rangle \leq \Phi_1^*(\ell), \forall \ell \in C_c']. \tag{3.62}$$

Young's inequality $\langle u, \ell\rangle \leq \Phi(u) + \Phi^*(\ell)$ and $u \in \Gamma \Leftrightarrow \Phi(u) = 0$ for all u, ℓ give the direct implication. For the converse, choosing $\ell = \varepsilon_{(a,b)}$ in the right-hand side of (3.62), one obtains for all $(a, b) \in \mathcal{S}$, $u(a, b) \leq \Phi^*(\varepsilon_{(a,b)})$. But, $\Phi^*(\varepsilon_{(a,b)}) = c(a, b)$ by Proposition 3.12-d. This proves (3.62) and completes the proof of the lemma. $\square$

3.6. **Optimal plan : completing the proofs of Theorem 3.3 and 3.5.** We are now in position to complete the proofs of these results.

Proof of Theorem 3.3. • Proof of (1). *Sufficient condition.* Let $\pi \in P(\mu, \nu)$ be such that $\int_{A\times B} c\, d\pi < \infty$. Let φ and ψ satisfy (3.27). Because of Lemma 3.26, one obtains that $\operatorname{ls}\varphi$ and $\operatorname{ls}\psi$ still satisfy (3.27) as well as $\operatorname{ls}\varphi \in L_1(A, \mu)$ and $\operatorname{ls}\psi \in L_1(B, \nu)$. Thanks to Lemma 3.31, there exists some $\omega \in \overline{\Upsilon}$ (see (3.28)) such that $\langle \omega, (\mu, \nu)\rangle = \int_{A\times B} c\, d\pi$. But, with Lemma 3.29-c: $T_2^*\omega \in \overline{\Gamma}$. Therefore, one can apply Theorem 3.24-b which insures that π is optimal.

Necessary condition. Let π be an optimal plan. Because of Theorem 3.24-b there exists $\omega \in \mathcal{X}_1^*$ such that $\eta := T_2^*\omega \in \overline{\Gamma}$ and (3.21-b) holds. With Lemma 3.43, one sees that $\widetilde{\eta}$ defined by (3.44) satisfies $\widetilde{\eta} \leq c$ on $\mathcal{S}$ and $\widetilde{\eta} = c$ on $\operatorname{supp}\pi \bigcap \{c < \infty\}$. By (3.16), for all $(a, b) \in \mathcal{S}$ we have $\widetilde{\eta}(a, b) = \widetilde{\omega}_A(a) + \widetilde{\omega}_B(b)$ where $\widetilde{\omega}_A(a) = \langle \omega_A, \epsilon_a\rangle$ and $\widetilde{\omega}_B(b) = \langle \omega_B, \epsilon_b\rangle$. One concludes the proof, taking $\varphi = \mathbf{1}_{\mathcal{S}_A}\widetilde{\omega}_A$ and $\psi = \mathbf{1}_{\mathcal{S}_B}\widetilde{\omega}_B$ where $\mathcal{S}_A$ and $\mathcal{S}_B$ are the canonical projections of $\mathcal{S}$ on A and B.

• Proof of (2). It appears from Lemmas 3.31 and 3.43 that the optimal functions (φ, ψ) and the optimal linear form $\omega \in \mathcal{X}_1^*$ associated with π by the KKT condition (3.21), see Theorem A.8, are related to each other by

$$\omega(\epsilon_a, \epsilon_b) = \varphi \oplus \psi(a, b), \text{ for } \pi\text{-a.e. } (a, b) \in A\times B. \tag{3.63}$$

Therefore, (3.21) and (3.4) express the same KKT condition. If φ and ψ are measurable, then they are integrable by Lemma 3.26-1 and they solve $(\overline{K})$ by Theorem 3.24. This proves statement (a). In the general situation (b), replacing (φ, ψ) by $(\operatorname{ls}\varphi, \operatorname{ls}\psi)$, one concludes similarly by means of Lemma 3.26-2. $\square$

Proof of Theorem 3.5. By Theorem 3.3 there exist functions φ_1 and ψ_1 satisfying (3.27). By Lemma 3.26 there exist functions φ_2 and ψ_2 such that $\varphi_2 \in L_1(A, \mu)$ and $\psi_2 \in L_1(B, \nu)$. Now with Lemma 3.31, one can extend (φ_2, ψ_2) in the sense of (3.63) into $\omega \in \mathcal{X}_1^*$ such that $\omega(\kappa) \leq \Lambda^*(|\kappa|)$, $\forall \kappa \in \mathcal{X}_1$. But, this is clearly equivalent to $|\omega(\kappa)| \leq \Lambda^*(|\kappa|)$, $\forall \kappa \in \mathcal{X}_1$. Applying Lemma 3.43 and taking $\varphi = \mathbf{1}_{\mathcal{S}_A}\widetilde{\omega}_A$ and $\psi = \mathbf{1}_{\mathcal{S}_B}\widetilde{\omega}_B$ as in the proof of the necessary condition of Theorem 3.3 leads to the desired result. $\square$

4. The proofs of the results of Section 2

The results of Section 2 are a summing up of Proposition 4.7, Lemma 4.11 , Proposition 4.12, Corollary 4.17, Lemma 4.19, Proposition 4.20, Proposition 4.30 and Proposition 4.38.

We are going to apply the general results of the Lagrangian approach to the minimization problem (P) which are recalled at Appendix A. We use the notations of Appendix A.

4.1. **Preliminary technical results.** Recall that $|u|_\Phi = \inf\{\alpha > 0; \Phi_\pm(u/\alpha) \le 1\}$ with $\Phi_\pm(u) = \max(\Phi(u), \Phi(-u))$. Its associated dual uniform norm is

$$|\ell|^*_\Phi \overset{\Delta}{=} \sup_{u, |u|_\Phi \le 1} |\langle u, \ell\rangle|, \quad \ell \in \mathcal{L}_1$$

on $\mathcal{L}_1$. The topological dual space of $(\mathcal{L}_1, |\cdot|^*_\Phi)$ is denoted by $\mathcal{L}_1'$. It is the topological bidual space of $(\mathcal{U}_1, |\cdot|_\Phi)$.
Similarly, recall that $|y|_\Lambda = \inf\{\alpha > 0; \Lambda_\pm(y/\alpha) \le 1\}$ with $\Lambda_\pm(y) = \max(\Lambda(y), \Lambda(-y))$. Its associated dual uniform norm is

$$|x|^*_\Lambda \overset{\Delta}{=} \sup_{y, |y|_\Lambda \le 1} |\langle y, x\rangle|, \quad x \in \mathcal{X}_1$$

on $\mathcal{X}_1$. The topological dual space of $(\mathcal{X}_1, |\cdot|^*_\Lambda)$ is denoted by $\mathcal{X}_1'$. It is the topological bidual space of $(\mathcal{Y}_1, |\cdot|_\Lambda)$.
The adjoint operator $T_1^\sharp$ which appears at Lemma 4.1-f below is defined as follows. For all $\omega \in \mathcal{X}_1'$ and all $\ell \in \mathcal{L}_1$, $|\langle T_1^\sharp \omega, \ell\rangle_{\mathcal{L}_1^*, \mathcal{L}_1}| = |\langle \omega, T\ell\rangle_{\mathcal{X}_1', \mathcal{X}_1}|$

Lemma 4.1. *Let us assume* (H_Φ) *and* (H_T).
(a) $\operatorname{dom}\Phi^* \subset \mathcal{L}_1$ *and* $\operatorname{dom}\Lambda^* \subset \mathcal{X}_1$
(b) $T(\operatorname{dom}\Phi^*) \subset \operatorname{dom}\Lambda^*$ *and* $T\mathcal{L}_1 \subset \mathcal{X}_1$
(c) T *is* $\sigma(\mathcal{L},\mathcal{U})$-$\sigma(\mathcal{X},\mathcal{Y})$-*continuous*
(d) $T_2^* : \mathcal{X}_1^* \to \mathcal{L}_1^*$ *is* $\sigma(\mathcal{X}_1^*, \mathcal{X}_1)$-$\sigma(\mathcal{L}_1^*, \mathcal{L}_1)$-*continuous*
(e) $T_1 : \mathcal{L}_1 \to \mathcal{X}_1$ *is* $|\cdot|^*_\Phi$-$|\cdot|^*_\Lambda$-*continuous*
(f) $T_1^\sharp \mathcal{X}_1' \subset \mathcal{L}_1'$
(g) $T_1^*\mathcal{Y}_1 \subset \mathcal{U}_1$ *and* $T_1^* : \mathcal{Y}_1 \to \mathcal{U}_1$ *is* $\sigma(\mathcal{Y}_1, \mathcal{X}_1)$-$\sigma(\mathcal{U}_1, \mathcal{L}_1)$-*continuous*
(h) $T_1 : \mathcal{L}_1 \to \mathcal{X}_1$ *is* $\sigma(\mathcal{L}_1,\mathcal{U}_1)$-$\sigma(\mathcal{X}_1, \mathcal{Y}_1)$-*continuous*

Proof. • Proof of (a). For all $\ell \in \mathcal{L}$ and $\alpha > 0$, Young's inequality yields: $\langle u, \ell\rangle = \alpha\langle \ell, u/\alpha\rangle \le [\Phi(u/\alpha)+\Phi^*(\ell)]\alpha$, for all $u \in \mathcal{U}$. Hence, for any $\alpha > |u|_\Phi$, $\langle u, \ell\rangle \le [1+\Phi^*(\ell)]\alpha$. It follows that $\langle u, \ell\rangle \le [1 + \Phi^*(\ell)]|u|_\Phi$. Considering $-u$ instead of u, one gets

$$|\langle u, \ell\rangle| \le [1 + \Phi^*(\ell)]|u|_\Phi, \forall u \in \mathcal{U}, \ell \in \mathcal{L}. \tag{4.2}$$

It follows that $\operatorname{dom}\Phi^* \subset \mathcal{L}_1$. One proves $\operatorname{dom}\Lambda^* \subset \mathcal{X}_1$ similarly.

• Proof of (b). Let us consider $|\cdot|_{\Phi^*_\pm}$ and $|\cdot|_{\Lambda^*_\pm}$ the gauge functionals of the level sets $\{\Phi^*_\pm \le 1\}$ and $\{\Lambda^*_\pm \le 1\}$. It is easy to show that

$$\Lambda^*_\pm(x) \le \Phi^*_\pm(\ell), \text{ for all } \ell \in \mathcal{L} \text{ and } x \in \mathcal{X} \text{ such that } T\ell = x \tag{4.3}$$

Therefore, $T(\operatorname{dom}\Phi^*_\pm) \subset \operatorname{dom}\Lambda^*_\pm$. On the other hand, by Proposition B.1 (see the Appendix), the linear space spanned by $\operatorname{dom}\Phi^*_\pm$ is $\operatorname{dom}|\cdot|_{\Phi^*_\pm}$ and the linear space spanned by $\operatorname{dom}\Lambda^*_\pm$ is $\operatorname{dom}|\cdot|_{\Lambda^*_\pm}$. But, $\operatorname{dom}|\cdot|_{\Phi^*_\pm} = \operatorname{dom}|\cdot|^*_\Phi = \mathcal{L}_1$ and $\operatorname{dom}|\cdot|_{\Lambda^*_\pm} = \operatorname{dom}|\cdot|^*_\Lambda = \mathcal{X}_1$ by Proposition B.1 again. Hence, $T\mathcal{L}_1 \subset \mathcal{X}_1$.

• Proof of (c). To prove that T is continuous, one has to show that for any $y \in \mathcal{Y}$, $\ell \in \mathcal{L} \mapsto \langle y, T\ell\rangle \in \mathbb{R}$ is continuous. We get $\ell \mapsto \langle y, T\ell\rangle = \langle T^*y, \ell\rangle$ which is continuous since (H_{T1}) is $T^*y \in \mathcal{U}$.

• Proof of (d). It is a direct consequence of $T\mathcal{L}_1 \subset \mathcal{X}_1$. See the proof of (c).

• Proof of (e). We know by Proposition B.1 that $|\cdot|_{\Phi^*_\pm} \sim |\cdot|^*_\Phi$ and $|\cdot|_{\Lambda^*_\pm} \sim |\cdot|^*_\Lambda$ are equivalent norms on $\mathcal{L}_1$ and $\mathcal{X}_1$ respectively. For all $\ell \in \mathcal{L}_1$, $|T\ell|^*_\Lambda \le 2|T\ell|_{\Lambda^*_\pm} = 2\inf\{\alpha > 0; \Lambda^*_\pm(T\ell/\alpha) \le 1\} \le 2\inf\{\alpha > 0; \Phi^*_\pm(\ell/\alpha) \le 1\}$. This last inequality follows

from (4.3). Going on, we get $|T\ell|^*_\Lambda \le 2|\ell|_{\Phi^*_\pm} \le 4|\ell|^*_\Phi$, which proves that T_1 shares the desired continuity property with $\|T_1\| \le 4$.

• Proof of (f). Let us take $\omega \in \mathcal{X}_1'$. For all $\ell \in \mathcal{L}_1$, $|\langle T_1^\sharp \omega, \ell\rangle_{\mathcal{L}_1^*,\mathcal{L}_1}| = |\langle \omega, T\ell\rangle_{\mathcal{X}_1',\mathcal{X}_1}| \le \|\omega\|_{\mathcal{X}_1'} |T\ell|^*_\Lambda \le \|\omega\|_{\mathcal{X}_1'}\|T_1\||\ell|^*_\Phi$ where $\|T_1\| < \infty$, thanks to (e). Hence, $T_1^\sharp\omega$ stands in $\mathcal{L}_1'$.

• Proof of (g). Let us take $y \in \mathcal{Y}_1$. We've just seen that $T_1^* y$ stands in $\mathcal{L}_1'$. Let us show that in addition, it is the strong limit of a sequence in $\mathcal{U}$. Indeed, there exists a sequence (y_n) in $\mathcal{Y}$ such that $\lim_{n\to\infty} y_n = y$ in $(\mathcal{Y}_1, |\cdot|_\Lambda)$. Hence, for all $\ell \in \mathcal{L}_1$, $|\langle T_1^* y_n - T_1^* y, \ell\rangle_{\mathcal{L}_1^*,\mathcal{L}_1}| = |\langle y_n - y, T\ell\rangle_{\mathcal{Y}_1,\mathcal{X}_1}| \le \|T_1\||y_n - y|_\Lambda |\ell|^*_\Phi$ and $\sup_{\ell\in\mathcal{L}_1, |\ell|^*_\Phi\le 1} |\langle T_1^* y_n - T_1^* y, \ell\rangle| \le \|T_1\||y_n - y|_\Lambda$ tends to 0 as n tends to infinity, where $T_1^* y_n = T^* y_n$ belongs to $\mathcal{U}$ for all $n \ge 1$ by (H_{T1}). Consequently, $T_1^* y \in \mathcal{U}_1$.
The continuity statement now follows from (d).

• Proof of (h).

By (b), T_1 maps $\mathcal{L}_1$ into $\mathcal{X}_1$ and because of (g): $T^*\mathcal{Y}_1 \subset \mathcal{U}_1$. Hence, for all $y \in \mathcal{Y}_1$, $\ell \mapsto \langle T_1\ell, y\rangle_{\mathcal{X}_1,\mathcal{Y}_1} = \langle \ell, T^* y\rangle_{\mathcal{L}_1,\mathcal{U}_1}$ is $\sigma(\mathcal{L}_1,\mathcal{U}_1)$-continuous. This completes the proof of Lemma 4.1. □

Let Φ_0^*, Λ_0^* and Λ_1^* be the convex conjugates of Φ_0, Λ_0 and Λ_1 for the dual pairings $\langle \mathcal{U}, \mathcal{L}\rangle$, $\langle \mathcal{Y}, \mathcal{X}\rangle$ and $\langle \mathcal{Y}_1, \mathcal{X}_1\rangle$.

Lemma 4.4. *Under the hypotheses (H_Φ) and (H_T), we have*

(a) $\Phi_0 = \Phi_1 \le \Phi$ *on* $\mathcal{U}$ *(a')* $\Lambda_0 = \Lambda_1 \le \Lambda$ *on* $\mathcal{Y}$
(b) $\Phi^* = \Phi_0^*$ *on* $\mathcal{L}$ *(b')* $\Lambda^* \le \Lambda_0^*$ *on* $\mathcal{X}$
(c) $\Phi^* = \Phi_0^* = \Phi_1^*$ *on* $\mathcal{L}_1$ *(c')* $\Lambda^* \le \Lambda_0^* \le \Lambda_1^*$ *on* $\mathcal{X}_1$

Proof. (a) follows directly from Lemma 4.1-a, (a') from (a) and (b') from (a').
(b) follows from the general fact that the convex conjugates of a function and its convex lower semicontinuous regularization match.
Let us show (c). As $\mathcal{U}$ is a dense subspace of $\mathcal{U}_1$, we obtain that the restriction of Φ^* to $\mathcal{L}_1$ is also the convex conjugate of Φ (restricted to $\mathcal{L}_1$) for the dual pairing $\langle \mathcal{U}_1, \mathcal{L}_1\rangle$. Now, with the same argument as in (b), this implies that $\Phi^* = \Phi_1^*$ on $\mathcal{L}_1$.
(c') follows from (a'), the fact that $\mathcal{Y}$ is a dense subset of $\mathcal{Y}_1$, the weak continuity of T_1^* which is proved at Lemma 4.1-g and the lower semicontinuity of Φ_1. □

Lemma 4.5. *Under the hypothesis (H_Φ),*

(a) $\Phi^* = \Phi_0^*$ *is* $\sigma(\mathcal{L},\mathcal{U})$*-inf-compact and*
(b) Φ_1^* *is* $\sigma(\mathcal{L}_1,\mathcal{U}_1)$*-inf-compact.*

Proof. • Proof of (b). We first prove that Φ_1^* is $\sigma(\mathcal{L}_1,\mathcal{U}_1)$-inf-compact. Recall that we already obtained at (4.2) that $|\langle u, \ell\rangle| \le [1 + \Phi^*(\ell)]|u|_\Phi$, for all $u \in \mathcal{U}$ and $\ell \in \mathcal{L}$. By completion, one deduces that for all $\ell \in \mathcal{L}_1$ and $u \in \mathcal{U}_1$, $|\langle u, \ell\rangle| \le [1 + \Phi_1^*(\ell)]|u|_\Phi$ (recall that $\Phi^* = \Phi_1^*$ on $\mathcal{L}_1$, Lemma 4.4-c.) Hence, $\Phi_1^*(\ell) \le A$ implies that $|\ell|^*_\Phi \le A+1$. Therefore, the level set $\{\Phi_1^* \le A\}$ is relatively $\sigma(\mathcal{L}_1,\mathcal{U}_1)$-compact.
By construction, Φ_1^* is $\sigma(\mathcal{L}_1,\mathcal{U}_1)$-lower semicontinuous. Hence, $\{\Phi_1^* \le A\}$ is $\sigma(\mathcal{L}_1,\mathcal{U}_1)$-closed and $\sigma(\mathcal{L}_1,\mathcal{U}_1)$-compact.

• Proof of (a). As $\Phi^* = \Phi_0^* = \Phi_1^*$ on $\mathcal{L}_1$ (Lemma 4.4-c), $\operatorname{dom}\Phi^* \subset \mathcal{L}_1$ (Lemma 4.1-a) and $\mathcal{U} \subset \mathcal{U}_1$, it follows from the $\sigma(\mathcal{L}_1,\mathcal{U}_1)$-inf-compactness of Φ_1^* that $\Phi^* = \Phi_0^*$ is $\sigma(\mathcal{L},\mathcal{U})$-inf-compact. □

4.2. **A first dual equality.** In this section we only consider the basic spaces $\mathcal{U}, \mathcal{L}, \mathcal{Y}$ and $\mathcal{X}$. Let us begin applying Appendix A with $\langle P, A\rangle = \langle \mathcal{U}, \mathcal{L}\rangle$ and $\langle B, Q\rangle = \langle \mathcal{Y}, \mathcal{X}\rangle$ and the topologies are the weak topologies $\sigma(\mathcal{L},\mathcal{U})$, $\sigma(\mathcal{U},\mathcal{L})$, $\sigma(\mathcal{X},\mathcal{Y})$ and $\sigma(\mathcal{Y},\mathcal{X})$. The function to be minimized is $f(\ell) = \Phi^*(\ell)+\delta_C(T\ell)$, $\ell \in \mathcal{L}$ where $\delta_C(x) = \begin{cases} 0 & \text{if } x \in C \\ +\infty & \text{if } x \notin C \end{cases}$ denotes the convex indicator of C. The perturbation F of f is Fenchel's one:

$$F_0(\ell, x) = \Phi^*(\ell) + \delta_C(T\ell + x), \ell \in \mathcal{L}, x \in \mathcal{X}.$$

We assume (H_{T1}): $T^*\mathcal{Y} \subset \mathcal{U}$, so that the duality diagram is

$$\begin{array}{ccc} \Big\langle \mathcal{U} & , & \mathcal{L} \Big\rangle \\ {\scriptstyle T^*}\big\uparrow & & \big\downarrow{\scriptstyle T} \\ \Big\langle \mathcal{Y} & , & \mathcal{X} \Big\rangle \end{array} \qquad \text{(Diagram 0)}$$

The analogue of F for the dual problem is

$$G_0(y, u) \overset{\triangle}{=} \inf_{\ell,x}\{\langle y, x\rangle - \langle u, \ell\rangle + F_0(\ell, x)\} = \inf_{x\in C}\langle y, x\rangle - \Phi_0(T^*y + u).$$

The corresponding value functions are

$$\begin{aligned} \varphi_0(x) &= \inf\{\Phi_0^*(\ell); \ell \in \mathcal{L} : T\ell \in C - x\}, \quad x \in \mathcal{X} \\ \gamma_0(u) &= \sup_{y\in\mathcal{Y}}\{\inf_{x\in C}\langle y, x\rangle - \Phi_0(T^*y + u)\}, \quad u \in \mathcal{U}. \end{aligned}$$

The primal and dual problems are (P) and (D_0).

Lemma 4.6. *Assuming (H_Φ) and (H_{T1}), if C is a $\sigma(\mathcal{X},\mathcal{Y})$-closed convex set, F_0 is jointly closed convex on $\mathcal{L} \times \mathcal{X}$.*

Proof. As T is linear continuous (Lemma 4.1-c) and C is closed convex, $\{(\ell, x); T\ell+x \in C\}$ is closed convex in $\mathcal{L} \times \mathcal{X}$. As Φ^* is closed convex on $\mathcal{L}$, its epigraph is closed convex in $\mathcal{L} \times \mathbb{R}$. It follows that $\operatorname{epi} F_0 = (\mathcal{X} \times \operatorname{epi} \Phi^*) \cap \{(\ell, x); T\ell + x \in C\}$ is closed convex, which implies that F_0 is convex and lower semicontinuous. As it is nowhere equal to $-\infty$ (since $\inf F_0 \geq \inf \Phi^* > -\infty$, F_0 is also a closed convex function. □

Therefore, assuming that C is a $\sigma(\mathcal{X},\mathcal{Y})$-closed convex set, one can apply the general theory of Appendix A since the perturbation function F_0 satisfies the assumptions (A.1) and (A.3).

Proposition 4.7. *Let us assume that (H_Φ) and (H_T) hold. If C is convex and $\sigma(\mathcal{X},\mathcal{Y})$-closed, we have the dual equality*

$$\inf(P) = \sup(D_0) \in [0, \infty]. \tag{4.8}$$

In particular, for all x in $\mathcal{X}$, we have the little dual equality

$$\inf\{\Phi^*(\ell); \ell \in \mathcal{L}, T\ell = x\} = \Lambda_0^*(x) \in [0, \infty]. \tag{4.9}$$

Proof. The identity (4.9) is a special case of (4.8) with $C = \{x\}$.
To prove (4.8), we consider separately the cases where $\inf(P) < +\infty$ and $\inf(P) = +\infty$.

Case where $\inf(P) < +\infty$. Thanks to Theorem A.6-b', it is enough to prove that γ_0 is upper semicontinuous at $u = 0$. We are going to prove that γ_0 is continuous at $u = 0$. Indeed, for all $u \in \mathcal{U}$,

$$-\gamma_0(u) = \inf_y\{\Phi_0(T^*y + u) - \inf_{x\in C}\langle y, x\rangle\} \leq \Phi_0(u) \leq \Phi(u)$$

where the first inequality is obtained taking $y=0$. The norm $|\cdot|_\Phi$ is designed so that Φ_0 is bounded above on a $|\cdot|_\Phi$-neighbourhood of zero. By the previous inequality, so is the convex function $-\gamma_0$. Therefore, $-\gamma_0$ is $|\cdot|_\Phi$-continuous on $\operatorname{icordom}(-\gamma_0)\ni 0$. As it is convex and $\mathcal{L}_1=(\mathcal{U},|\cdot|_\Phi)'$, it is also $\sigma(\mathcal{U},\mathcal{L}_1)$-lower semicontinuous and a fortiori $\sigma(\mathcal{U},\mathcal{L})$-lower semicontinuous , since $\mathcal{L}_1\subset\mathcal{L}$.

Case where $\inf(P)=+\infty$. Note that $\sup(D_0)\ge -\Phi_0(0)=0>-\infty$, so that we can apply Theorem A.6-b. It is enough to prove that

$$\operatorname{ls}\varphi_0(0)=+\infty$$

in the situation where $\varphi_0(0)=\inf(P)=+\infty$. We have $\operatorname{ls}\varphi_0(0)=\sup_{V\in\mathcal{N}(0)}\inf\{\Phi_0^*(\ell);\ell: T\ell\in C+V\}$ where $\mathcal{N}(0)$ is the set of all the $\sigma(\mathcal{X},\mathcal{Y})$-open neighbourhoods of $0\in\mathcal{X}$. It follows that for all $V\in\mathcal{N}(0)$, there exists $\ell\in\mathcal{L}$ such that $T\ell\in C+V$ and $\Phi_0^*(\ell)\le\operatorname{ls}\varphi_0(0)$. This implies that

$$T(\{\Phi_0^*\le\operatorname{ls}\varphi_0(0)\})\cap(C+V)\neq\emptyset,\quad\forall V\in\mathcal{N}(0).\tag{4.10}$$

On the other hand, $\inf(P)=+\infty$ is equivalent to: $T(\operatorname{dom}\Phi_0^*)\cap C=\emptyset$.
Now, we prove ad absurdum that $\operatorname{ls}\varphi_0(0)=+\infty$. Suppose that $\operatorname{ls}\varphi_0(0)<+\infty$. Because of $T(\operatorname{dom}\Phi_0^*)\cap C=\emptyset$, we have a fortiori

$$T(\{\Phi_0^*\le\operatorname{ls}\varphi_0(0)\})\cap C=\emptyset.$$

As Φ_0^* is inf-compact (Lemma 4.5-a) and T is weakly continuous (Lemma 4.1-c), $T(\{\Phi_0^*\le \operatorname{ls}\varphi_0(0)\})$ is a $\sigma(\mathcal{X},\mathcal{Y})$-compact subset of $\mathcal{X}$. Clearly, it is also convex. But C is assumed to be closed and convex, so that by Hahn-Banach theorem, C and $T(\{\Phi_0^*\le\operatorname{ls}\varphi_0(0)\})$ are *strictly* separated. This contradicts (4.10), considering open neighbourhoods V of the origin in (4.10) which are open half-spaces. Consequently, $\operatorname{ls}\varphi_0(0)=+\infty$. This completes the proof of the proposition. □

4.3. **Primal attainment and dual equality.** We are going to consider the following duality diagram, see Section 2.3:

$$\begin{array}{ccc}\Big\langle\ \mathcal{U}_1 & , & \mathcal{L}_1\ \Big\rangle\\ T_1^*\Big\uparrow & & \Big\downarrow T_1\\ \Big\langle\ \mathcal{Y}_1 & , & \mathcal{X}_1\ \Big\rangle\end{array}\tag{Diagram 1}$$

Note that the inclusions $T_1\mathcal{L}_1\subset\mathcal{X}_1$ and $T_1^*\mathcal{Y}_1\subset\mathcal{U}_1$ which are stated in Lemma 4.1 are necessary to validate this diagram.
Let F_1,G_1 and γ_1 be the analogous functions to F_0, G_0 and γ_0. Denoting φ_1 the primal value function, we obtain

$$\begin{aligned}F_1(\ell,x)&=\Phi_1^*(\ell)+\delta_{C_1}(T_1\ell+x),\ell\in\mathcal{L}_1,\quad x\in\mathcal{X}_1\\ G_1(y,u)&=\inf_{x\in C_1}\langle y,x\rangle-\Phi_1(T_1^*y+u),\quad y\in\mathcal{Y}_1,u\in\mathcal{U}_1\quad\text{and}\\ \varphi_1(x)&=\inf\{\Phi_1^*(\ell);\ell\in\mathcal{L}_1:T_1\ell\in C_1-x\},\quad x\in\mathcal{X}_1,\\ \gamma_1(u)&=\sup_{y\in\mathcal{Y}_1}\{\inf_{x\in C_1}\langle y,x\rangle-\Phi_1(T_1^*y+u)\},\quad u\in\mathcal{U}_1.\end{aligned}$$

It appears that the primal and dual problems are (P_1) and (D_1).

Lemma 4.11. *Assuming* (H_Φ) *and* (H_T), *the problems* (P) *and* (P_1) *are equivalent: they have the same solutions and* $\inf(P)=\inf(P_1)\in[0,\infty]$.

Proof. It is a direct consequence of $\operatorname{dom}\Phi^* \subset \mathcal{L}_1$, $T\mathcal{L}_1 \subset \mathcal{X}_1$ and $\Phi^* = \Phi_1^*$ on $\mathcal{L}_1$, see Lemma 4.1-a,b and Lemma 4.4-c. □

Proposition 4.12 (Primal attainment and dual equality)**.** *Assume that* (H_Φ) *and* (H_T) *hold.*

(a) *For all* x *in* $\mathcal{X}_1$, *we have the little dual equality*

$$\inf\{\Phi^*(\ell); \ell \in \mathcal{L}, T\ell = x\} = \Lambda_1^*(x) \in [0, \infty]. \tag{4.13}$$

Assume that in addition (H_C) *holds.*

(b) *We have the dual equalities*

$$\inf(P) = \sup(D_1) \in [0, \infty] \tag{4.14}$$

$$\inf(P) = \inf(P_1) = \inf_{x \in C_1} \Lambda_1^*(x) \in [0, \infty] \tag{4.15}$$

(c) *If in addition* $\inf(P) < \infty$, *then* (P) *is attained in* $\mathcal{L}_1$.

(d) *Let* $\bar{\ell} \in \mathcal{L}_1$ *be a solution to* (P), *then* $\bar{x} \triangleq T\bar{\ell}$ *is a solution to* $(P_{1,\mathcal{X}})$ *and* $\inf(P) = \Phi^*(\bar{\ell}) = \Lambda_1^*(\bar{x})$.

Proof. • We begin with the proof of (4.14). As, $\inf(P) = \inf(P_1)$ by Lemma 4.11, we have to show that $\inf(P_1) = \sup(D_1)$. We consider separately the cases where $\inf(P_1) < +\infty$ and $\inf(P_1) = +\infty$.

Case where $\inf(P_1) < +\infty$. Because of (H_C), F_1 is jointly convex and $F_1(\ell, \cdot)$ is $\sigma(\mathcal{X}_1, \mathcal{Y}_1)$-closed convex for all $\ell \in \mathcal{L}_1$. As $T_1^*\mathcal{Y}_1 \subset \mathcal{U}_1$ (Lemma 4.1), one can apply the approach of Appendix A to the duality Diagram 1. Therefore, by Theorem A.6-b', the dual equality holds if γ_1 is $\sigma(\mathcal{U}_1, \mathcal{L}_1)$-upper semicontinuous at 0. As in the proof of Proposition 4.7, we have $-\gamma_1(u) \le \Phi_1(u)$, for all $u \in \mathcal{U}_1$. But Φ_1 is the $\sigma(\mathcal{U}_1, \mathcal{L}_1)$-lower semicontinuous regularization of $\tilde{\Phi}(u) = \begin{cases} \Phi(u) & \text{if } u \in \mathcal{U} \\ +\infty & \text{otherwise} \end{cases}$, $u \in \mathcal{U}_1$ and Φ is bounded above by 1 on the ball $\{u \in \mathcal{U}; |u|_\Phi < 1\}$. As $\mathcal{L}_1 = (\mathcal{U}_1, |\cdot|_\Phi)'$, Φ_1 is also the $|\cdot|_\Phi$-regularization of $\tilde{\Phi}$. Therefore, Φ_1 is bounded above by 1 on $\{u \in \mathcal{U}_1; |u|_\Phi < 1\}$, since $\{u \in \mathcal{U}; |u|_\Phi < 1\}$ is $|\cdot|_\Phi$-dense in $\{u \in \mathcal{U}_1; |u|_\Phi < 1\}$. As $-\gamma_1 (\le \Phi_1)$ is convex and bounded above on a $|\cdot|_\Phi$-neighbourhood of 0, it is $|\cdot|_\Phi$-continuous on $\operatorname{icordom}(-\gamma_1) \ni 0$. Hence, it is $\sigma(\mathcal{U}_1, \mathcal{L}_1)$-lower semicontinuous at 0.

Case where $\inf(P_1) = +\infty$. This proof is a transcription of the second part of the proof of Proposition 4.7, replacing T by T_1, C by C_1, all the subscripts 0 by 1 and using the preliminary results: Φ_1^* is inf-compact (Lemma 4.5) and T_1 is weakly continuous (Lemma 4.1-h). This completes the proof of (4.14).

• The identity (4.13) is simply (4.14) with $C_1 = \{x\}$.

• Let us prove (c). By Lemma 4.1-h, T_1 is $\sigma(\mathcal{L}_1, \mathcal{U}_1)$-$\sigma(\mathcal{X}_1, \mathcal{Y}_1)$-continuous. Since C_1 is $\sigma(\mathcal{X}_1, \mathcal{Y}_1)$-closed, $\{\ell \in \mathcal{L}_1; T\ell \in C_1\}$ is $\sigma(\mathcal{L}_1, \mathcal{U}_1)$-closed. As Φ_1^* is $\sigma(\mathcal{L}_1, \mathcal{U}_1)$-inf-compact (Lemma 4.5), it achieves its infimum on the closed set $\{\ell \in \mathcal{L}_1; T\ell \in C_1\}$ if $\inf(P_1) = \inf(P) < \infty$.

• Let us prove (4.15). The dual equality (4.14) gives us, for all $x_o \in C_1$, $\inf(P_1) = \sup_{y \in \mathcal{Y}_1}\{\inf_{x \in C_1}\langle y, x\rangle - \Lambda_1(y)\} \le \sup_{y \in \mathcal{Y}_1}\{\langle x_o, y\rangle - \Lambda_1(y)\} = \Lambda_1^*(x_o)$. Therefore

$$\inf(P_1) \le \inf_{x \in C_1} \Lambda_1^*(x). \tag{4.16}$$

In particular, equality holds instead of inequality if $\inf(P_1) = +\infty$. Suppose now that $\inf(P_1) < \infty$. From statement (c), we already know that there exists $\bar{\ell} \in \mathcal{L}_1$ such that $\bar{x} \triangleq T\bar{\ell} \in C_1$ and $\inf(P_1) = \Phi^*(\bar{\ell})$. Clearly $\inf(P_1) \le \inf\{\Phi^*(\ell); T\ell = \bar{x}, \ell \in \mathcal{L}_1\} \le \Phi^*(\bar{\ell})$.

Hence, $\inf(P_1) = \inf\{\Phi_1^*(\ell); T\ell = \bar{x}, \ell \in \mathcal{L}_1\}$. By the little dual equality (4.13) we have $\inf\{\Phi_1^*(\ell); T\ell = \bar{x}, \ell \in \mathcal{L}_1\} = \Lambda_1^*(\bar{x})$. Finally, we have obtained $\inf(P_1) = \Lambda_1^*(\bar{x})$ with $\bar{x} \in C_1$. Together with (4.16), this leads us to the desired identity: $\inf(P_1) = \inf_{x \in C_1} \Lambda_1^*(x)$.

- Finally, (d) is a by-product of the proof of (4.15). □

The following result is an improvement of Lemma 4.4-c'.

Corollary 4.17. *We have* $\operatorname{dom}\Lambda_1^* \subset \operatorname{dom}\Lambda_0^*$, $\operatorname{dom}\Lambda_1^* \subset \mathcal{X}_1$ *and in restriction to* $\mathcal{X}_1$, $\Lambda_0^* = \Lambda_1^*$.

Proof. The first part is already proved at Lemma 4.1-a. The matching $\Lambda_0^* = \Lambda_1^*$ follows from (4.9) and (4.13). □

4.4. **Dual attainment.** We now consider the following duality diagram

$$\begin{array}{ccc} \Big\langle\ \mathcal{L}_1 & , & \mathcal{L}_1^*\ \Big\rangle \\ T_1 \Big\downarrow & & \Big\uparrow T_2^* \\ \Big\langle\ \mathcal{X}_1 & , & \mathcal{X}_1^*\ \Big\rangle \end{array} \tag{Diagram 2}$$

where the topologies are the respective weak topologies. The associated perturbation functions are

$$\begin{aligned} F_2(\ell, x) &= \Phi_1^*(\ell) + \delta_{C_1}(T\ell + x), \quad \ell \in \mathcal{L}_1, x \in \mathcal{X}_1 \\ G_2(\zeta, \omega) &= \inf_{x \in C_1} \langle x, \omega \rangle - \Phi_2(T_2^*\omega + \zeta), \quad \zeta \in \mathcal{L}_1^*, \omega \in \mathcal{X}_1^* \end{aligned}$$

As $F_2 = F_1$, the primal problem is (P_1) and its value function is φ_1 :

$$\varphi_1(x) = \inf_{x' \in C_1 - x} \Lambda_1^*(x'), \quad x \in \mathcal{X}_1 \tag{4.18}$$

where we used (4.13). The dual problem is (D_2).

Assume that $\inf(P) < \infty$. We know by Proposition 4.12-d that $(P_{1,\mathcal{X}})$ admits at least a solution $\bar{x} = T\bar{\ell}$ where $\bar{\ell}$ is a solution to (P_1). Let us consider the following new minimization problem

$$\text{minimize } \Phi_1^*(\ell) \quad \text{subject to} \quad T\ell = \bar{x}, \quad \ell \in \mathcal{L}_1 \tag{$P_1^{\bar{x}}$}$$

Of course $\bar{\ell}$ is a solution to (P_1) if and only if it is a solution to $(P_1^{\bar{x}})$ where $\bar{x} = T\bar{\ell}$. Since our aim is to derive a representation formula for $\bar{\ell}$, it is enough to build our duality schema upon $(P_1^{\bar{x}})$ rather than upon (P_1). The associated perturbation functions are

$$\begin{aligned} F_2^{\bar{x}}(\ell, x) &= \Phi_1^*(\ell) + \delta_{\{\bar{x}\}}(T\ell + x), \quad \ell \in \mathcal{L}_1, x \in \mathcal{X}_1 \\ G_2^{\bar{x}}(\zeta, \omega) &= \langle \bar{x}, \omega \rangle - \Phi_2(T_2^*\omega + \zeta), \quad \zeta \in \mathcal{L}_1^*, \omega \in \mathcal{X}_1^* \end{aligned}$$

As $F_2^{\bar{x}}$ is F_1 with $C_1 = \{\bar{x}\}$, the primal problem is $(P_1^{\bar{x}})$ and its value function is

$$\varphi_1^{\bar{x}}(x) = \Lambda_1^*(\bar{x} - x), \quad x \in \mathcal{X}_1.$$

The dual problem is

$$\text{maximize} \quad \langle \bar{x}, \omega \rangle - \Lambda_2(\omega), \quad \omega \in \mathcal{X}_1^* \tag{$D_2^{\bar{x}}$}$$

Lemma 4.19. *Under the hypotheses* (H_Φ) *and* (H_T), Λ_1^* *is* $\sigma(\mathcal{X}_1, \mathcal{Y}_1)$*-inf-compact.*

Proof. By (4.13): $\inf\{\Phi_1^*(\ell); \ell \in \mathcal{L}_1, T_1\ell = x\} = \Lambda_1^*(x)$ for all $x \in \mathcal{X}_1$ (note that $\Phi^* = \Phi_1^*$ on $\mathcal{L}_1$ by Lemma 4.4-c.) As T_1 is continuous (Lemma 4.1-h) and Φ_1^* is inf-compact(Lemma 4.5), it follows that Λ_1^* is also inf-compact. □

Proposition 4.20 (Dual attainment)**.** *Assume that* (H_Φ), (H_T) *and* (H_C) *hold.*

(a) *Suppose that*

$$C \cap \operatorname{icordom} \Lambda_1^* \neq \emptyset. \tag{4.21}$$

Then the dual problem (D_2) *is attained in* $\mathcal{X}_1^*$.

(b) *Suppose that* $C \cap \operatorname{dom} \Lambda_1^* \neq \emptyset$. *Then,* $\inf(P) < \infty$ *and we know (see Proposition 4.12-d) that* $(P_{1,\mathcal{X}})$ *admits at least a solution. If in addition, there exists a solution* $\bar{x}$ *to* $(P_{1,\mathcal{X}})$ *such that*

$$\bar{x} \in \operatorname{diffdom} \Lambda_1^*, \tag{4.22}$$

then the dual problem $(D_2^{\bar{x}})$ *is attained in* $\mathcal{X}_1^*$.

Proof. • Proof of (a). As $F_2 = F_1$, one can apply the approach of Appendix A to the duality Diagram 2. Let us denote φ_1^{**} the $\sigma(\mathcal{X}_1, \mathcal{Y}_1)$-lower semicontinuous regularization of φ_1 and φ_2^{**} its $\sigma(\mathcal{X}_1, \mathcal{X}_1^*)$-lower semicontinuous regularization. Since $\mathcal{X}_1$ separates $\mathcal{Y}_1$, the inclusion $\mathcal{Y}_1 \subset \mathcal{X}_1^*$ holds. It follows that $\varphi_1^{**}(0) \leq \varphi_2^{**}(0) \leq \varphi_1(0)$. But we have (4.14) which is $\varphi_1^{**}(0) = \varphi_1(0)$. Therefore, one also obtains $\varphi_2^{**}(0) = \varphi_1(0)$ which is the dual equality

$$\inf(P_1) = \sup(D_2) \tag{4.23}$$

and one can apply Theorem A.6-c which gives

$$\operatorname{argmax}(D_2) = -\partial\varphi_1(0). \tag{4.24}$$

It remains to show that the value function φ_1 given at (4.18) is such that

$$\partial\varphi_1(0) \neq \emptyset. \tag{4.25}$$

As the considered dual pairing $\langle \mathcal{X}_1, \mathcal{X}_1^* \rangle$ is the saturated algebraic pairing, for (4.25) to be satisfied, by the geometric version of Hahn-Banach theorem, it is enough that $0 \in \operatorname{icordom} \varphi_1$. But this holds provided that the constraint qualification (4.21) is satisfied.

• Proof of (b). Let us specialize to the special case where $C_1 = \{\bar{x}\}$. The dual equality (4.23) becomes

$$\inf(P_1^{\bar{x}}) = \sup(D_2^{\bar{x}}) \tag{4.26}$$

and (4.25) becomes $\partial\varphi_1^{\bar{x}}(0) \neq \emptyset$ which is directly implied by (4.22). □

Remark 4.27. The dual equality (4.26) is

$$\Lambda_1^* = \Lambda_2^* \tag{4.28}$$

where these convex conjugates are to be taken respectively with respect to $\langle \mathcal{X}_1, \mathcal{Y}_1 \rangle$ and $\langle \mathcal{X}_1, \mathcal{X}_1^* \rangle$. Denoting $\overline{\Lambda}_1$ and $\overline{\Lambda}_2$ the convex $\sigma(\mathcal{X}_1^*, \mathcal{X}_1)$-lower semicontinuous regularizations of Λ_1 and Λ_2, (4.28) implies the identity

$$\overline{\Lambda}_1 = \overline{\Lambda}_2. \tag{4.29}$$

Usual results about convex conjugation tell us that $\Lambda_1^*(\bar{x}) = \sup_{\omega \in \mathcal{X}_1^*} \{\langle \bar{x}, \omega \rangle - \overline{\Lambda}_1(\omega)\} = \sup(D_2^{\bar{x}})$ and the above supremum is attained at $\bar{\omega}$ if and only if $\bar{\omega} \in \partial_{\mathcal{X}_1^*} \Lambda_1^*(\bar{x})$. This is the attainment statement in Proposition 4.20-b.

4.5. **Dual representation of the minimizers.** We keep the framework of Diagram 2 and derive the KKT relations in this situation. The Lagrangian associated with $F_2 = F_1$ and Diagram 2 is for any $\ell \in \mathcal{L}_1, \omega \in \mathcal{X}_1^*$,

$$\begin{aligned} K_2(\ell, \omega) &\triangleq \inf_{x \in \mathcal{X}_1} \{\langle x, \omega \rangle + \Phi_1^*(\ell) + \delta_{C_1}(T\ell + x)\}, \\ &= \Phi_1^*(\ell) - \langle T\ell, \omega \rangle + \inf_{x \in C_1} \langle x, \omega \rangle. \end{aligned}$$

Proposition 4.30. *Assume that* (H_Φ), (H_T) *and* (H_C) *hold.*

(a) *Any $\bar\ell \in \mathcal{L}_1$ is a solution to (P_1) if and only if there exist some $\bar\omega \in \mathcal{X}_1^*$ such that the following three statements hold*
   (1) $T\bar\ell \in C$
   (2) $\langle T\bar\ell, \bar\omega\rangle \le \langle x, \bar\omega\rangle$ for all $x \in C_1$
   (3) and the following representation formula holds
$$\bar\ell \in \partial_{\mathcal{L}_1}\Phi_2(T_2^*\bar\omega) \tag{4.31}$$
   More, these three statements hold if and only if: $\bar\ell$ is solution to (P_1), $\bar\omega$ is a solution to (D_2) and $\inf(P_1) = \sup(D_2)$.
   Statement (4.31) is equivalent to the Young's identity
$$\Phi^*(\bar\ell) + \Phi_2(T_2^*\bar\omega) = \langle T\bar\ell, \bar\omega\rangle. \tag{4.32}$$
(b) *(Assumption (H_C) is useless here). Any $\bar\ell \in \mathcal{L}_1$ is a solution to $(P^{\bar x})$ if and only if $T\bar\ell = \bar x$ and there exists some $\bar\omega \in \mathcal{X}_1^*$ such that (4.31) or equivalently (4.32) holds.*
   More, this occurs if and only if: $\bar\ell$ is a solution to (P), $\bar\omega$ is a solution to $(D_2^{\bar x})$ with $\bar x := T\bar\ell$ and $\inf(P^{\bar x}) = \sup(D_2^{\bar x})$.

Proof. This proof is an application of Theorem A.8. Under the general assumptions (H_Φ), (H_T) and (H_C), we have seen at Proposition 4.20 that the dual equalities (4.23) and (4.26) hold true. In both situations (a) and (b), $(\bar\ell, \bar\omega)$ is a saddle-point; all we have to do is to translate the KKT relations (A.10) and (A.11).

• Proof of (a). With K_2 as above, (A.10) and (A.11) are $\partial_\ell K_2(\bar\ell, \bar\omega) \ni 0$ and $\partial_\omega(-K_2)(\bar\ell, \bar\omega) \ni 0$. Since $-\langle T\ell, \omega\rangle$ is locally weakly upper bounded as a function of ω around $\bar\omega$ and as a function of ℓ around $\bar\ell$, one can apply (Rockafellar, [7], Theorem 20) to derive $\partial_\ell K_2(\bar\ell, \bar\omega) = \partial\Phi^*(\bar\ell) - T_2^*\bar\omega$ and $\partial_\omega(-K_2)(\bar\ell, \bar\omega) = \partial(-\inf_{x\in C_1}\langle x, \cdot\rangle) + T\bar\ell$. Therefore the KKT relations are
$$\begin{aligned} T_2^*\bar\omega &\in \partial\Phi^*(\bar\ell) &(4.33)\\ -T\bar\ell &\in \partial(\delta^*_{-C_1})(\bar\omega) &(4.34)\end{aligned}$$
where $\delta^*_{-C_1}$ is the convex conjugate of the convex indicator of $-C_1$.
As a convex conjugate, Φ^* is a closed convex functions. Its convex conjugate is Φ_2. Therefore (4.33) is equivalent to the following equivalent statements
$$\begin{aligned}&\bar\ell \in \partial\Phi_2(T_2^*\bar\omega)\\ &\Phi^*(\bar\ell) + \Phi_2(T_2^*\bar\omega) = \langle \bar\ell, T_2^*\bar\omega\rangle\end{aligned}$$
Similarly, as a convex conjugate $\delta^*_{-C_1}$ is a closed convex functions. Its convex conjugate is $\delta_{-\bar C_1}$ where $\bar C_1$ stands for the $\sigma(\mathcal{X}_1, \mathcal{X}_1^*)$-closure of C_1. Of course, as C_1 is $\sigma(\mathcal{X}_1, \mathcal{Y}_1)$-closed by hypothesis (H_C), it is a fortiori $\sigma(\mathcal{X}_1, \mathcal{X}_1^*)$-closed, so that $\bar C_1 = C_1$. Therefore (4.34) is equivalent to
$$\delta_{C_1}(T\bar\ell) + \delta^*_{-C_1}(\bar\omega) = \langle -T\bar\ell, \bar\omega\rangle. \tag{4.35}$$
It follows from (4.35) that $\delta_{C_1}(T\bar\ell) < \infty$ which is equivalent to $T\bar\ell \in C_1$.
Now (4.35) is $-\langle T\bar\ell, \bar\omega\rangle = \delta^*_{-C_1}(\bar\omega) = -\inf_{x\in C_1}\langle x, \bar\omega\rangle$ which is $\langle T\bar\ell, \bar\omega\rangle = \inf_{x\in C_1}\langle x, \bar\omega\rangle$. This completes the proof of (a).

• Proof of (b). This follows directly from (a) with $\bar x = T\bar\ell$ and $C_1 = \{\bar x\}$. □

Remark 4.36. Thanks to Proposition 4.12-d, (4.32) leads us to
$$\Lambda_1^*(\bar x) + \Lambda_2(\bar\omega) = \langle \bar x, \bar\omega\rangle \tag{4.37}$$
for all $\bar x \in \operatorname{dom}\Lambda_1^*$ and all $\bar\omega \in \mathcal{X}_1^*$ solution to $(D_2^{\bar x})$. By Young's inequality: $\Lambda_2^*(\bar x) + \overline{\Lambda}_2(\bar\omega) \ge \langle \bar x, \bar\omega\rangle$ and the identities (4.28), (4.37), we see that $\overline{\Lambda}_2(\bar\omega) \ge \Lambda_2(\bar\omega)$. But, the

reversed inequality always holds true. Therefore, we have $\overline{\Lambda}_2(\bar\omega) = \Lambda_2(\bar\omega)$. This proves that $\Lambda_2 = \overline{\Lambda}_2$ on $\mathrm{dom}\,\Lambda_2$.

Proposition 4.38. *Assume that (H_Φ), (H_T) and (H_C) hold. Any solution $\bar\omega$ of (D_2) or $(D_2^{\bar x})$ shares the following properties*

(a) *$\bar\omega$ stands in the $\sigma(\mathcal{X}_1^*, \mathcal{X}_1)$-closure of* $\mathrm{dom}\,\Lambda_1$.

(b) *$T_2^*\bar\omega$ stands in the $\sigma(\mathcal{L}_1^*, \mathcal{L}_1)$-closures of $T_1^*(\mathrm{dom}\,\Lambda_1)$ and* $\mathrm{dom}\,\Phi$.

(c) *For any $x_o \in \mathcal{X}_1$, let us denote $j_{D_{x_o}}$ and $j_{-D_{x_o}}$ the gauge functionals on $\mathcal{X}_1$ of the convex sets D_{x_o} and $-D_{x_o}$ where $D_{x_o} = \{x \in \mathcal{X}_1; \Lambda_1^*(x_o + x) \le \Lambda_1^*(x_o) + 1\}$.*
   - *Let $\bar\omega$ be any solution of (D_2). Then, for any x_o in $C \cap \mathrm{icordom}\,\Lambda_1^*$, $\bar\omega$ is $j_{D_{x_o}}$-upper semicontinuous and $j_{-D_{x_o}}$-lower semicontinuous at* 0.
   - *Let $\bar\omega$ be any solution of $(D_2^{\bar x})$ with $\bar x \in \mathrm{icordom}\,\Lambda_1^*$. Then, $\bar\omega$ is $j_{D_{\bar x}}$-upper semicontinuous and $j_{-D_{\bar x}}$-lower semicontinuous at* 0.

Proof. • Proof of (a). Because of (4.37), we have $\bar\omega \in \mathrm{dom}\,\Lambda_2$. As $\overline{\Lambda}_2 \le \Lambda_2$ and $\overline{\Lambda}_1 = \overline{\Lambda}_2$ (see (4.29)), we obtain $\bar\omega \in \mathrm{dom}\,\overline{\Lambda}_1$ which implies (a).

• Proof of (b). It follows from (a) and the continuity of T_2^*, see Lemma 4.1-d that $T_2^*\bar\omega$ is in the $\sigma(\mathcal{L}_1^*, \mathcal{L}_1)$-closure of $T_1^*(\mathrm{dom}\,\Lambda_1)$. On the hand, $T_2^*\bar\omega \in \mathrm{dom}\,\Phi_2$ and Φ_2 is the $\sigma(\mathcal{L}_1^*, \mathcal{L}_1)$-closed convex closure of Φ. It follows that $T_2^*\bar\omega$ is in the $\sigma(\mathcal{L}_1^*, \mathcal{L}_1)$-closure of $\mathrm{dom}\,\Phi$.

• Proof of (c). Let $\bar\omega \in \mathrm{argmax}(D_2)$. By (4.18) and (4.24), for all $x \in \mathcal{X}_1$ and any $x_o \in C_1$, $\langle -\bar\omega, x\rangle \le \varphi_1(x) - \varphi_1(0) \le \Lambda_1^*(x_o - x) - \varphi(0) \le \Lambda_1^*(x_o - x)$. It follows that $\langle \bar\omega, x\rangle \le \Lambda_1^*(x_o) + 1$ for all $x \in D_{x_o}$. This implies that for all $x \in \mathcal{X}_1$, $\langle \bar\omega, x\rangle \le [1 + \Lambda_1^*(x_o)] j_{D_{x_o}}(x)$. Since $j_D(-x) = j_{-D}(x)$, we finally obtain

$$-[1 + \Lambda_1^*(x_o)] j_{-D_{x_o}}(x) \le \langle \bar\omega, x\rangle \le [1 + \Lambda_1^*(x_o)] j_{D_{x_o}}(x), \forall x \in \mathcal{X}_1$$

for any $x_o \in C_1$, which is the desired result. Choosing x_o in $C_1 \cap \mathrm{icordom}\,\Lambda_1^*$ implies that $j_{D_{x_o}}$ is a nondegerate homogeneous functional.
The second case where $\bar\omega \in \mathrm{argmax}(D_2^{\bar x})$ is a specialization of the previous one. □

Appendix A. A short reminder about convex minimization

To quote easily and precisely some well-known results of convex minimization while proving our abstract results at Section 4, we give a short overview of the approach to convex minimization problems by means of conjugate duality as developed in Rockafellar's monograph [7]. For complete proofs of these results, one can also have a look at the author's lecture notes [3].

Let A be a vector space and $f : A \to [-\infty, +\infty]$ an extended real convex function. We consider the following convex minimization problem

$$\text{minimize } f(a), a \in A \tag{$\mathcal{P}$}$$

Let Q be another vector space. The perturbation of the objective function f is a function $F : A \times Q \to [-\infty, +\infty]$ such that for $q = 0 \in Q$, $F(\cdot, 0) = f(\cdot)$. The problem $(\mathcal{P})$ is imbedded in a parametrized family of minimization problems

$$\text{minimize } F(a, q), a \in A \tag{$\mathcal{P}_q$}$$

The value function of $(\mathcal{P}_q)_{q\in Q}$ is

$$\varphi(q) \stackrel{\Delta}{=} \inf(\mathcal{P}_q) = \inf_{a\in A} F(a, q) \in [-\infty, +\infty], q \in Q.$$

Let us assume that the perturbation is chosen such that

$$F \text{ is jointly convex on } A \times Q. \tag{A.1}$$

Then, $(\mathcal{P}_q)_{q\in Q}$ is a family of convex minimization problems and the value function φ is convex.

Let B be a vector space in dual pairing with Q. This means that B and Q are locally convex topological vector spaces in separating duality such that their topological dual spaces B' and Q' satisfy $B' = Q$ and $Q' = B$ up to some isomorphisms. The Lagrangian associated with the perturbation F and the duality $\langle B, Q\rangle$ is

$$K(a,b) \triangleq \inf_{q\in Q}\{\langle b,q\rangle + F(a,q)\}, a \in A, b \in B. \tag{A.2}$$

Under (A.1), K is a convex-concave function. Assuming in addition that F is chosen such that

$$q \mapsto F(a,q) \text{ is a closed convex function for any } a \in A, \tag{A.3}$$

one can reverse the conjugate duality relation (A.2) to obtain

$$F(a,q) = \sup_{b\in B}\{K(a,b) - \langle b,q\rangle\}, \forall a \in A, q \in Q \tag{A.4}$$

Introducing another vector space P in separating duality with A we define the function

$$G(b,p) \triangleq \inf_{a\in A}\{K(a,b) - \langle a,p\rangle\}, b \in B, p \in P. \tag{A.5}$$

This formula is analogous to (A.4). Going on symmetrically, one interprets G as the concave perturbation of the objective concave function

$$g(b) \triangleq G(b,0), b \in B$$

associated with the concave maximization problem

$$\textit{maximize } g(b), b \in B \tag{$\mathcal{D}$}$$

which is the dual problem of $(\mathcal{P})$. It is imbedded in the family of concave maximization problems $(\mathcal{D}_p)_{p\in P}$

$$\textit{maximize } G(b,p), b \in B \tag{$\mathcal{D}_p$}$$

whose value function is

$$\gamma(p) \triangleq \sup_{b\in B} G(b,p), p \in P.$$

Since G is jointly concave, γ is also concave. We have the following diagram

$$\begin{array}{ccccccc}
 & & \gamma(p) & & f(a) & & \\
 & \Big\langle & P & , & A & \Big\rangle & \\
G(b,p) & & & K(a,b) & & & F(a,q) \\
 & \Big\langle & B & , & Q & \Big\rangle & \\
 & & g(b) & & \varphi(q) & &
\end{array}$$

The concave conjugate of the function f with respect to the dual pairing $\langle Y, X\rangle$ is $f^{\hat{*}}(y) = \inf_x\{\langle y,x\rangle - f(x)\}$ and its superdifferential at x is $\widehat{\partial} f(x) = \{y \in Y; f(x') \le f(x) + \langle y, x'-x\rangle\}$.

Theorem A.6. *We assume that $\langle P, A\rangle$ and $\langle B, Q\rangle$ are topological dual pairings.*

(a) *We have* $\sup(\mathcal{D}) = \varphi^{**}(0)$. *Hence, the dual equality* $\inf(\mathcal{P}) = \sup(\mathcal{D})$ *holds if and only if* $\varphi(0) = \varphi^{**}(0)$.

(b) *In particular,*

$$\left.\begin{array}{l}\bullet\ F \textit{ is jointly convex}\\ \bullet\ \varphi \textit{ is lower semicontinuous at } 0\\ \bullet\ \sup(\mathcal{D}) > -\infty\end{array}\right\} \Rightarrow \inf(\mathcal{P}) = \sup(\mathcal{D})$$

(c) *If the dual equality holds, then*

$$\operatorname{argmax} g = -\partial\varphi(0).$$

Let us assume in addition that (A.1) and (A.3) are satisfied.

(a') *We have* $\inf(\mathcal{P}) = \gamma^{\hat{*}\hat{*}}(0)$. *Hence, the dual equality* $\inf(\mathcal{P}) = \sup(\mathcal{D})$ *holds if and only if* $\gamma(0) = \gamma^{\hat{*}\hat{*}}(0)$.

(b') *In particular,*

$$\left.\begin{array}{l}\bullet\ \gamma \textit{ is upper semicontinuous at } 0\\ \bullet\ \inf(\mathcal{P}) < +\infty\end{array}\right\} \Rightarrow \inf(\mathcal{P}) = \sup(\mathcal{D})$$

(c') *If the dual equality holds, then*

$$\operatorname{argmin} f = -\widehat{\partial}\gamma(0).$$

Definition A.7 (Saddle-point)**.** *One says that* $(\bar{a}, \bar{b}) \in A \times B$ *is a* saddle-point *of the function* K *if*

$$K(\bar{a}, b) \leq K(\bar{a}, \bar{b}) \leq K(a, \bar{b}), \quad \forall a \in A, b \in B.$$

Theorem A.8 (Saddle-point theorem and KKT relations)**.** *The following statements are equivalent.*

(1) The point $(\bar{a}, \bar{b})$ *is a saddle-point of the Lagrangian* K

(2) $f(\bar{a}) \leq g(\bar{b})$

(3) The following three statements hold

(a) we have the dual equality*:* $\sup(\mathcal{D}) = \inf(\mathcal{P})$,

(b) $\bar{a}$ *is a solution to the primal problem* $(\mathcal{P})$ *and*

(c) $\bar{b}$ *is a solution to the dual problem* $(\mathcal{D})$.

In this situation, one also gets

$$\sup(\mathcal{D}) = \inf(\mathcal{P}) = K(\bar{a}, \bar{b}) = f(\bar{a}) = g(\bar{b}). \tag{A.9}$$

Moreover, $(\bar{a}, \bar{b})$ *is a saddle-point of* K *if and only if it satisfies*

$$\partial_a K(\bar{a}, \bar{b}) \ni 0 \tag{A.10}$$

$$\widehat{\partial}_b K(\bar{a}, \bar{b}) \ni 0 \tag{A.11}$$

where the subscript a *or* b *indicates the unfixed variable.*

Appendix B. Gauge functionals associated with a convex function

The following result is well-known, but since I didn't find a reference for it, I give its short proof.

Let $\theta : S \to [0, \infty]$ be an extended positif convex function on a vector space S, such that $\theta(0) = 0$. Let S^* be the algebraic dual space of S and θ^* the convex conjugate of $\theta :$

$$\theta^*(r) \overset{\triangle}{=} \sup_{s \in S}\{\langle r, s\rangle - \theta(s)\}, r \in S^*.$$

It is easy to show that $\theta^* : S^* \to [0,\infty]$ and $\theta^*(0) = 0$. We denote $C_\theta \overset{\triangle}{=} \{\theta \le 1\}$ and $C_{\theta^*} \overset{\triangle}{=} \{\theta^* \le 1\}$ the unit level sets of θ and θ^*. The gauge functionals to be considered are

$$\begin{aligned} j_\theta(s) &\overset{\triangle}{=} \inf\{\alpha > 0; s \in \alpha C_\theta\} = \inf\{\alpha > 0; \theta(s/\alpha) \le 1\} \in [0,\infty], s \in S. \\ j_{\theta^*}(r) &\overset{\triangle}{=} \inf\{\alpha > 0; r \in \alpha C_{\theta^*}\} = \inf\{\alpha > 0; \theta^*(r/\alpha) \le 1\} \in [0,\infty], r \in S^*. \end{aligned}$$

As 0 belongs to C_θ and C_{θ^*}, one easily proves that j_θ and j_{θ^*} are positively homogeneous. Similarly, as C_θ and C_{θ^*} are convex sets, j_θ and j_{θ^*} are convex functions.

Proposition B.1. *Let $\theta : S \to [0,\infty]$ be an extended positif convex function on a vector space S, such that $\theta(0) = 0$ as above. Then for all $r \in S^*$, we have*

$$\frac{1}{2} j_{\theta^*}(r) \le \delta^*_{C_\theta}(r) \overset{\triangle}{=} \sup_{s\in C_\theta} \langle r, s\rangle \le 2 j_{\theta^*}(r).$$

We also have

$$\operatorname{cone}\operatorname{dom}\theta^* = \operatorname{dom} j_{\theta^*} = \operatorname{dom}\delta^*_{C_\theta}$$

where $\operatorname{cone}\operatorname{dom}\theta^*$ *is the convex cone (with vertex 0) generated by* $\operatorname{dom}\theta^*$.

Proof. • Let us first show that $\delta^*_{C_\theta}(r) \le 2 j_{\theta^*}(r)$ for all $r \in S^*$. If $j_{\theta^*}(r) > 0$, then for all $s \in C_\theta$, $\langle r, s\rangle = \langle r/j_{\theta^*}(r), s\rangle j_{\theta^*}(r) \le [\theta(s) + \theta^*(r/j_{\theta^*}(r))] j_{\theta^*}(r) \le (1+1) j_{\theta^*}(r)$.
If $j_{\theta^*}(r) = 0$, then $\theta^*(tr) \le 1$ for all $t > 0$. For any $s \in C_\theta$, we get $\langle r, s\rangle = \frac{1}{t}\langle tr, s\rangle \le \frac{1}{t}[\theta(s) + \theta^*(tr)] \le 2/t$. Letting t tend to infinity, one obtains that $\langle r, s\rangle \le 0$.
• Let us show that $j_{\theta^*}(r) \le 2\delta^*_{C_\theta}(r)$. If $\delta^*_{C_\theta}(r) = \infty$, there is nothing to prove. So, let us suppose that $\delta^*_{C_\theta}(r) < \infty$. As $0 \in C_\theta$, we have $\delta^*_{C_\theta}(r) \ge 0$.
First case: $\delta^*_{C_\theta}(r) > 0$. For all $s \in S$ and $\epsilon > 0$, we have $s/[j_\theta(s) + \epsilon] \in C_\theta$. It follows that $\langle r/\delta^*_{C_\theta}(r), s\rangle = \langle r, s/[j_\theta(s)+\epsilon]\rangle \frac{j_\theta(s)+\epsilon}{\delta^*_{C_\theta}(r)} \le \delta^*_{C_\theta}(r) \frac{j_\theta(s)+\epsilon}{\delta^*_{C_\theta}(r)} = j_\theta(s) + \epsilon$. Therefore, $\langle r/\delta^*_{C_\theta}(r), s\rangle \le j_\theta(s)$, for all $s \in S$.
If s doesn't belong to C_θ, then $j_\theta(s) \le \theta(s)$. This follows from the the assumptions on θ : convex function such that $\theta(0) = 0 = \min\theta$ and the positive homogeneity of j_θ. Otherwise, if s belongs to C_θ, we have $j_\theta(s) \le 1$. Hence, $\langle r/\delta^*_{C_\theta}(r), s\rangle \le \max(1, \theta(s)), \forall s \in S$. On the other hand, there exists $s_o \in S$ such that $\theta^*(r/[2\delta^*_{C_\theta}(r)]) \le \langle r/[2\delta^*_{C_\theta}(r)], s_o\rangle - \theta(s_o) + 1/2$. The last two inequalities provide us with $\theta^*(r/[2\delta^*_{C_\theta}(r)]) \le \frac{1}{2}\max(1, \theta(s_o)) - \theta(s_o) + \frac{1}{2} \le 1$ since $\theta(s_o) \ge 0$. We have proved that $j_{\theta^*}(r) \le 2\delta^*_{C_\theta}(r)$.
Second case: $\delta^*_{C_\theta}(r) = 0$. We have $\langle r, s\rangle \le 0$ for all $s \in C_\theta$. As $\operatorname{dom}\theta$ is a subset of the cone generated by C_θ, we also have for all $t > 0$ and $s \in \operatorname{dom}\theta$, $\langle tr, s\rangle \le 0$. Hence $\langle tr, s\rangle - \theta(s) \le 0$ for all $s \in S$ and $\theta^*(tr) \le 0$, for all $t \ge 0$. As $\theta^* \ge 0$, we have $\theta^*(tr) = 0$, for all $t \ge 0$. It follows that $j_{\theta^*}(r) = 0$. This completes the proof of the equivalence of j_{θ^*} and $\delta^*_{C_\theta}$.
• Finally, this equivalence implies that $\operatorname{dom} j_{\theta^*} = \operatorname{dom}\delta^*_{C_\theta}$ and as $\theta^*(0) = 0$ we have $0 \in \operatorname{dom}\theta^*$ which implies that $\operatorname{cone}\operatorname{dom}\theta^* = \operatorname{dom} j_{\theta^*}$. □

References

[1] L. Ambrosio and A. Pratelli. *Existence and stability results in the L^1-theory of optimal transportation. CIME Course*, volume 1813 of *Lecture Notes in Mathematics*, pages 123–160. Springer Verlag, 2003.

[2] J.M. Borwein and A.S. Lewis. Decomposition of multivariate functions. *Can. J. Math.*, 44(3):463–482, 1992.

[3] C. Léonard. A set of lecture notes on convex optimization with some applications to probability theory. Incomplete draft. Available online via `http://www.cmap.polytechnique.fr/~leonard/`.

[4] C. Léonard. Dominating points and entropic projections. Preprint, 2006.

[5] J. Neveu. *Bases mathématiques du calcul des probabilités.* Masson, Paris, 1970.

[6] S. Rachev and L. Rüschendorf. *Mass Transportation Problems. Vol I : Theory, Vol. II : Applications*. Probability and its applications. Springer Verlag, New York, 1998.

[7] R.T. Rockafellar. *Conjugate Duality and Optimization*, volume 16 of *Regional Conferences Series in Applied Mathematics*. SIAM, Philadelphia, 1974.

[8] L. Rüschendorf. On c-optimal random variables. *Statist. Probab. Lett.*, 27(3):267–270, 1996.

[9] W. Schachermayer and J. Teichman. Characterization of optimal transport plans for the Monge-Kantorovich problem. Preprint, 2006.

[10] C. Villani. *Topics in Optimal Transportation*. Graduate Studies in Mathematics 58. American Mathematical Society, Providence RI, 2003.

[11] C. Villani. Saint-Flour Lecture Notes. Optimal transport, old and new.
Available online via `http://www.umpa.ens-lyon.fr/~cvillani/`, 2005.

Modal-X, Université Paris 10 &

CMAP, École Polytechnique. 91128 Palaiseau Cedex, France
E-mail address: `christian.leonard@polytechnique.fr`